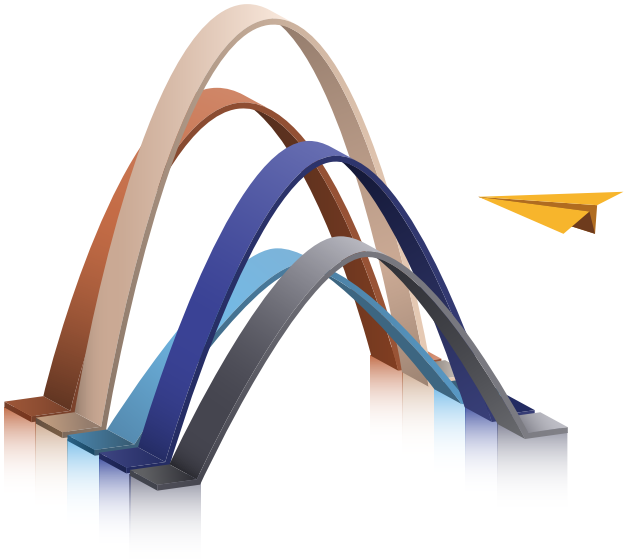
面向21世纪本科应用型经管规划教材
经济管理专业基础课系列



统计学

STATISTICS

刘徐方　王海文◎主　编
梁　旭　刘芳娜◎副主编

電子工業出版社
Publishing House of Electronics Industry
北京 • BEIJING

内 容 简 介

本书按照“应用型、技能型”人才培养目标，本着由浅入深、学以致用的原则，系统介绍统计学总论、统计调查与统计整理、综合指标、时间数列、统计指数、抽样推断、相关与回归分析、统计分析与统计报告等统计学基本知识，并通过指导学生实训实践培养其解决实际问题的能力。

本书具有知识系统、理论适中、案例经典、贴近实际、通俗易懂等特点，既可作为普通高等院校本科经济管理、物流、工商管理、国际贸易、电子商务等专业的首选教材，也可作为高职高专及成人高校的教材，还可用于统计工作者的在职培训，对于广大社会读者也是一本有益的参考读物。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目（CIP）数据

统计学 / 刘徐方，王海文主编. —北京：电子工业出版社，2017.9
面向 21 世纪本科应用型经管规划教材. 经济管理专业基础课系列
ISBN 978-7-121-32758-2

Ⅰ. ①统… Ⅱ. ①刘… ②王… Ⅲ. ①统计学－高等学校－教材 Ⅳ. ①C8

中国版本图书馆 CIP 数据核字（2017）第 232220 号

策划编辑：姜淑晶
责任编辑：张 京
印 刷：
装 订：
出版发行：电子工业出版社
北京市海淀区万寿路 173 信箱 邮编 100036
开 本：787×1092 1/16 印张：15.25 字数：390 千字
版 次：2017 年 9 月第 1 版
印 次：2017 年 9 月第 1 次印刷
定 价：46.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：（010）88254199，sjb@phei.com.cn。

序　言

Preface

随着我国改革开放的不断深入，我国经济已经连续30多年保持中高速稳定增长的态势。近年来，随着我国改革开放和社会主义市场经济的加速推进，随着国家“一带一路、互联互通”总体发展战略的制定和实施，我国经济迅速融入全球经济一体化的发展进程，中国市场国际化的特征越发凸显。

目前，我国正处于经济快速发展与社会变革的重要时期，随着经济转型、产业结构调整、传统企业改造，涌现出了一大批电子商务、文化创意、绿色生态及循环经济等新型产业。面对国际化市场的激烈竞争、面对新一轮的人才争夺，我国企业既要加快管理体制与运营模式的整改，也要注重加强经营理念与管理方法的不断创新，更加注重企业发展的本土化策略，抓紧网罗和培养具有创新意识和掌握新专业知识的技能型人才，这既是企业立于不败之地的根基，也是企业可持续发展的重要战略选择。

需求促进专业建设，市场驱动人才培养。为适应市场对经济管理专业人才多层次、多样化的需求，保证合理的人才结构，有必要开展多层次的经济管理技能培训与教育：一是加强学历教育；二是重视继续教育；三是开展有针对性的员工培训。

为了适应我国经济发展对“有思路、掌握技能、会操作、能应用”人才的需要，为了全面贯彻教育部关于“加强职业教育、强化实践、突出技能培养”的方针，针对我国高等教育“经济管理”专业知识老化、教材陈旧、重理论轻实践、缺乏实操技能训练的问题，根据企业用人与就业岗位的实际需要，结合应用型大学本科院校“经济管理”专业教学计划及课程设置与调整，我们组织北京工业大学、郑州大学、北京联合大学、青岛大学、北方工业大学、首钢工学院、大连工业大学、北京城市学院、燕山大学、北京财贸职业学院、吉林工程技术师范学院、北京朝阳职工大学、华北水利水电大学、北京西城社区学院、山东外贸职业学院等全国20多所院校经济管理专业的主讲教师和工商企业经理，共同精心编撰了此套教材，旨在迅速提高我国高等院校经济管理专业学生和工商企业从业者的专业技术素质，更好地服务于国家经济建设。

教材建设是应用型大学教育教学改革的重要组成部分，也是体现职业技能培养特色的关键。本套教材作为应用型大学经济管理专业的特色教材，融入了经济管理的最新实践教学理念、坚持以科学发展观为统领，力求严谨、注重与时俱进，根据国际经济发展的新形势和新特点，依照经济活动的基本过程和规律，全面贯彻国家“十二五”教育发展规划，按照企业对人的需求模式，加强实践能力训练，注重校企结合、贴近流通和工商企业业务实际，强化实践技能与岗位应用培养训练，并注重教学内容和教材结构的创新。

本套教材以专业基础课、专业课程为主，具体包括经济学基础、商务谈判与推销、统计学、基础会计、财政与金融、管理信息系统、公共关系与现代礼仪等教材。

由于本套教材紧密结合我国企业改革与经济发展、注重前瞻性，具有理论前沿性和实践操作性，注重实际应用和操作技能训练与培养，适应国家经济发展新常态的需要，对帮助学生尽快熟悉操作规程与业务管理、毕业后能够顺利走上社会就业具有特殊意义，因此本套教材既可作为应用型大学本科经济管理专业教学的首选教材，同时兼顾高职高专及成人高校的教学，也可以用于工商、通流、财贸等企业的在职员工培训。

牟惟仲

前　言

Forward

生产离不开统计，流通也离不开统计，随着国家经济建设的迅猛发展，无论是在制造业，还是在商务活动、经营领域，统计的作用日趋明显。在当今信息社会中，无论是科学研究，还是宏观调控、微观管理，乃至人们的社会生活，信息的获取与利用都非常重要，因而统计信息作为社会经济信息的主体越来越受到各行业、各领域人士的普遍关注。

统计学是高等教育经管类专业学生一门重要的专业基础核心课程。由于统计工作的政策性、专业性和应用性很强，因此已被广为应用在社会学、心理学、管理学、人口学、政治学、经济学及公共卫生等社会各个领域，并成为从业者的必备知识。

随着国家“一带一路、互联互通”总体发展战略的制定和实施，我国企业正在大步走出国门、积极参与经济建设。当前，面对全球经济一体化进程的加快和国际市场的激烈竞争，面对世界经济高度融合、社会经济生活节奏日趋快速化，对各类信息数据的需求明显加快和增多，其处理技术也更加复杂化和多样化。加强统计从业者应用技能培训、强化统计专业化综合业务素质培养，提高统计调查和统计计算与分析作业水平，这既是我国工商企事业单位长远发展的战略选择，也是本书出版的目的和意义。

本书作为高等教育与教学创新的特色教材，坚持以学科发展观为统领、广泛吸收国内外优秀统计教材的成果，注重基础知识体系完整，注重实践操作技能培养，注重教材内容创新，充分体现了职业性、实践性、应用性，不仅满足了为国家经济建设服务的社会需求，而且促进了我国统计知识普及和统计教育水平的提高。本书出版对帮助学生尽快熟悉统计操作规程、掌握业务岗位技能、毕业后顺利就业具有重要作用。

全书共 9 章，以学习者应用能力培养为主线，根据国家新修订并实施的《统计法》，依照统计工作与统计调查的基本过程和规律，结合“技能型、应用型”人才培养目标，具体介绍总论、统计调查、统计整理、综合指标、时间数列、统计指数、抽样推断、相关与回归分析、统计分析与统计报告等统计学基本知识，并通过指导学生实训实践培养解决实际问题的能力。

由于本书融入统计的最新实践教学理念，力求严谨，注重与时俱进，具有理论适中、知识系统、观点科学、案例经典、贴近实际、注重统计工作过程体验、突出职业教育等特点，因此本书既可以作为普通高等院校本科经济管理、物流、工商管理、国际贸易、电子商务等相关专业的首选教材，同时兼顾高职高专及成人高校的教学，也可以作为企事业统计从业者的岗位培训用书，并为广大社会创业人员提供有益的参考和帮助。

本书由李大军教授筹划并具体组织实施；刘徐方和王海文主编，刘徐方统稿；梁旭、刘

芳娜为副主编；本书由统计专家董力教授审定。作者编写分工：牟惟仲（序言），刘徐方（第1、2章），王海文（第3、4章），梁旭（第5、9章），刘芳娜（第6、7章），汤宁（第8章、各章实训习题）；华燕萍负责文字修改和版式调整，李晓新制作了教学课件。

在本书编写过程中，我们参阅、借鉴了大量国内外关于统计学的最新书刊资料并调查了业界的实践成果，得到了国家统计局、北京市统计局、调查咨询公司及有关专家的具体指导，在此一并表示感谢。为了方便教学，本书配有教学课件，读者可以从电子工业出版社网站（www.phei.com.cn）免费下载使用。因作者水平有限，书中难免有疏漏和不足之处，恳请专家、同行和广大读者批评指正。

编　者

目　录

Contents

第 1 章　总论 …… 1

学习目标要求 …… 1
重点难点提示 …… 1
教学引导案例 …… 1
第一节　统计学概述 …… 2
第二节　统计学的研究方法和统计工作过程 …… 8
第三节　统计学中的基本概念 …… 10
本章小结 …… 16
复习思考题 …… 16
实训基础训练 …… 17
课外延伸阅读 …… 18

第 2 章　统计调查 …… 21

学习目标要求 …… 21
重点难点提示 …… 21
教学引导案例 …… 21
第一节　统计调查概述 …… 23
第二节　统计调查方案 …… 26
第三节　统计调查方法 …… 30
第四节　统计调查报告 …… 40
本章小结 …… 46
复习思考题 …… 46
实训基础训练 …… 47
课外延伸阅读 …… 48

第 3 章　统计整理 …… 51

学习目标要求 …… 51
重点难点提示 …… 51
教学引导案例 …… 51
第一节　统计整理概述 …… 52
第二节　统计分组 …… 56
第三节　分配数列 …… 61
第四节　统计表和统计图 …… 68
本章小结 …… 74
复习思考题 …… 74
实训基础训练 …… 74
课外延伸阅读 …… 76

第 4 章　综合指标 …… 79

学习目标要求 …… 79
重点难点提示 …… 79
教学引导案例 …… 79
第一节　总量指标 …… 80
第二节　相对指标 …… 84
第三节　平均指标 …… 93
第四节　标志变异指标 …… 107
本章小结 …… 113
复习思考题 …… 113
实训基础训练 …… 114
课外延伸阅读 …… 115

第 5 章　时间数列 …… 117

学习目标要求 …… 117
重点难点提示 …… 117
教学引导案例 …… 117
第一节　时间数列概述 …… 119
第二节　时间数列的水平指标 …… 123
第三节　时间数列的速度指标 …… 131
本章小结 …… 137
复习思考题 …… 137
实训基础训练 …… 138
课外延伸阅读 …… 139

第 6 章 统计指数 …… 142

学习目标要求 …… 142
重点难点提示 …… 142
教学引导案例 …… 142
第一节 统计指数概述 …… 143
第二节 综合指数 …… 146
第三节 平均指数 …… 150
第四节 指数体系和因素分析 …… 153
第五节 指数在统计实践中的应用 …… 155
本章小结 …… 160
复习思考题 …… 160
实训基础训练 …… 161
课外延伸阅读 …… 162

第 7 章 抽样推断 …… 164

学习目标要求 …… 164
重点难点提示 …… 164
教学引导案例 …… 164
第一节 抽样推断概述 …… 165
第二节 抽样误差 …… 171
第三节 总体指标的估计方法 …… 176
第四节 样本容量确定与抽样方案设计 …… 180
本章小结 …… 185
复习思考题 …… 186
实训基础训练 …… 186
课外延伸阅读 …… 187

第 8 章 相关与回归分析 …… 189

学习目标要求 …… 189
重点难点提示 …… 189
教学引导案例 …… 189
第一节 相关分析概述 …… 190
第二节 简单线性相关分析 …… 194
第三节 回归分析 …… 198

本章小结 205
复习思考题 205
实训基础训练 206
课外延伸阅读 207

第 9 章　统计分析与统计报告 210

学习目标要求 210
重点难点提示 210
教学引导案例 210
第一节　统计分析概述 212
第二节　统计分析方法 216
第三节　统计分析报告 221
本章小结 227
复习思考题 228
实训基础训练 228
课外延伸阅读 228

参考文献 232



第1章 总 论



学习目标要求

1. 了解统计学的产生和发展过程，理解统计学的学科性质、含义与研究对象。
2. 熟悉统计研究的基本方法，掌握统计学中的基本概念。



重点难点提示

1. 统计总体和总体单位。
2. 指标、标志、变量、变量值等基本概念和作用，统计指标与数量标志的联系与区别。



教学引导案例

2016年7月社会消费品零售数据解读

国家统计局发布了2016年7月份社会消费品零售总额数据。对此，国家统计局贸易外经司高级统计师蔺涛进行了解读。

一、消费品零售继续保持两位数增长

7月份，社会消费品零售总额同比名义增长10.2%，增速虽比6月份回落0.4%，但比上半年仅低0.1%，与二季度持平；扣除价格因素后，实际增长9.8%（除特殊说明外，以下均为名义增长），增速比6月份低0.5%，但高于上半年0.1%；1~7月份社会消费品零售总额同比增长10.3%，增速与上半年持平，比上年同期仅低0.1%。总体来看，消费品市场运行态势平稳，继续保持在10%以上的较快增长区间。

二、基本生活类商品平稳增长，居住类商品增速加快

（1）基本生活类商品保持较快增长。7月份，限额以上单位吃穿用等基本生活类商品增长10.1%，比上月仅回落0.2%，增速继续保持在10%以上。其中，服装类增长9.4%，增速比上月加快1.9%；饮料类增长12.6%，比上月加快0.2%。

（2）居住类商品增速有所加快。7月份，限额以上单位建筑及装潢材料类、家具类同比分别增长15.0%和13.6%，增速分别比上月加快0.8%和0.2%。

三、网上零售业持续快速发展，消费升级类相关商品增长较快

一是网上商品零售额继续保持快速增长。1～7月份，全国实物商品网上零售额同比增长26.1%，增速虽比上半年回落0.5%，但仍高于社会消费品零售总额增速15.8%。

二是部分与消费升级相关商品增长较快。7月份，限额以上单位通信器材类、家用电器和音像器材类同比分别增长11.6%和11.5%；化妆品类增长9%，增速比上月加快1.1%；汽车类增长9.2%，增速连续两个月保持在9%以上，特别是新能源汽车、SUV、MPV，继续保持高速增长。

四、7月份名义增速回落主要是因为受与“618”电商促销相关商品零售回落等因素影响

近年来“618”电商年中大促已经超出平台、行业的范畴，成为线上线下、各行各业争相参与的消费大狂欢。今年“618”期间，有数十家来自线上线下的大型商家以“618”的名义参与促销，力度和规模超过往年。受上月“618”促销拉动消费提前等因素的影响，7月份网上零售、文化办公用品类、体育娱乐用品类商品零售增速回落幅度较大。据测算，7月份网上商品零售额增速比上月回落5%，影响社会消费品零售总额增速0.5%。

资料来源：国家统计局

http://www.stats.gov.cn/tjsj/sjjd/201608/t20160812_1387882.html

思考：

1. 该案例是2016年7月社会消费品零售总额数据资料的解读，从中可以得出哪些结论？

2. 这些数据是怎样得来的？如何保证这些数据资料的完整、准确和科学？

3. 统计对国民经济生活的意义是什么？

第一节　统计学概述

一、统计学的基本含义

“统计”一词有着丰富的内涵，一般讲到“统计”，可以从三个方面理解，即统计工作、统计资料和统计学。

1. 统计工作

统计工作又称统计实践，指利用科学的方法搜集、整理、分析和提供关于社会经济现象数量资料的工作的总称，是统计的基础。例如，各级统计部门对其所属地区的工业、农业及商业等方面的数据资料进行搜集、整理、分析等工作就是统计工作。

2. 统计资料

统计资料又称统计数据，指通过统计工作取得的、用来反映社会经济现象的数据资料的总称。统计工作所取得的各项数字资料及有关文字资料，大量的统计资料多以各种统计公报、统计年鉴、数据库及光盘等形式公布和收藏。

3. 统计学

统计学是指研究如何对统计资料进行搜集、整理和分析的理论与方法的科学。统计学既

是统计实践活动的经验总结和理论概括，又是指导统计工作的原理和原则。

统计的三种含义之间并不是相互孤立、相互排斥的，而是有一定的联系。它们的联系主要表现在两个方面。

（1）统计工作与统计资料

统计工作与统计资料是统计工作过程与活动成果的关系，统计工作的目的是取得统计资料。

（2）统计工作与统计学

统计工作与统计学是统计实践与理论的关系。一方面，统计学是统计工作的经验总结和理论概括；另一方面，统计学又指导统计工作的实践。

统计工作是先于统计学而发展起来的。自从有了国家以来，统计工作就随着社会经济的发展和国家管理的需要而发展起来，统计工作有四千多年的历史，而统计学只有 300 多年的历史。

统计工作、统计资料和统计学相互依存、相互联系，共同构成了一个完整的整体，这就是统计。

二、统计基本职能

《统计法》规定：“统计的基本任务是对国民经济和社会发展情况进行统计调查、统计分析，提供统计资料和统计咨询意见，实行统计监督。”这是国家基于统计工作特有的性质、功能和作用对统计职能和职责作出的界定。

统计基本职能通常概括为统计的信息职能、咨询职能和监督职能。

1. 信息职能

信息职能是统计的最基本职能，统计人员根据科学的统计指标体系和统计调查方法，搜集经济、科技和社会等各方面的信息，并向全社会提供各方面所需的信息，统计信息服务主要以统计报表或统计公报为载体。在当今社会，统计信息已成为企业经营活动、人们日常社会生活和国家进行宏观经济管理必不可少的重要依据。

2. 咨询职能

统计的咨询职能是统计信息职能的延续和深化。统计部门利用已掌握的丰富统计信息资源，运用科学的分析方法、先进的技术手段、通过全面综合分析和专题研究，为各级机关、团体、企业及社会公众提供可供选择的咨询建议和对策方案；统计咨询服务通常以统计咨询报告为载体。

3. 监督职能

统计的监督职能是指通过统计调查和分析，及时、准确地从总体上反映社会经济现象的运行状态，并对其进行全面、系统的检查、检测和预警，以供有关职能部门及时研究和解决发展中存在的问题，促使国民经济按照客观规律持续、快速和健康地发展。

在上述三个职能中，信息职能是统计职能的基础，咨询、监督都必须以统计信息为前提，统计咨询职能是统计信息职能的延续和深化，而统计监督职能则是统计信息、统计咨询职能的进一步拓展，并促进统计信息和统计咨询职能的优化。统计的信息职能、咨询职能和监督职能是一个有机整体，它们彼此依存、相辅相成。

三、统计学的研究对象

任何一门独立的学科都有区别于其他学科的独特研究对象。只有明确了研究对象，才能根据它的性质、特点采用相应的研究方法，达到认识现象客体规律性的目的。

统计学的研究对象是指统计研究所要认识的客体。它决定着统计科学的研究领域及相应的研究方法。一般来说，统计学的研究对象是客观事物的数量特征和数量关系。即把社会、自然现象总体的数量特征和数量关系作为自己的研究对象。统计学的研究对象具有如下特点。

1. 数量性

统计学研究对象的数量性，就是通过各种统计指标和指标体系来反映对象总体的规模、水平、速度、比例、效益和趋势等。可见，研究现象的数量问题是统计学研究对象的基本特征，数字是统计的语言。

客观事物的数量方面是指社会经济现象的规模、水平、结构、速度、比例关系、差别程度、发展速度、平均规模和水平、平均发展速度等。由于事物的质和量是密切联系的，因此，统计学在研究社会现象时，首先从定性研究开始，然后进行定量分析，最后认识社会现象的本质、特征或规律。

总之，统计学不是研究现象的纯数量关系，而是在质与量的辩证统一中研究现象的数量关系。现象的质和量是对立统一的两个方面，在研究现象的数量问题时，不能离开现象的质，应以现象质的分析为基础来确定现象数量表现的特定范围。

2. 总体性

由于社会经济现象错综复杂，个别现象所处的时间、地点和条件不同，表现出偶然性和不确定性，难以说明社会经济现象总体的本质和规律。只有以社会经济现象的总体为研究对象，即以构成总体的全部或足够多的单位为研究对象时，才能消除偶然性因素的影响，防止片面性，从而正确地揭示社会经济现象的本质和规律性。

统计认识社会经济现象的数量问题必须是对总体现象的认识，而非对个体现象的认识。例如，可以通过对一个国家或地区的众多工业企业的研究，了解工业企业的生产能力、生产规模、产品结构和工业品满足社会需要的程度等方面的情况。但如果只对该国或该地区的个别工业企业进行观察，就不可能得到整个工业产品的结构及其他重要信息，因为它不具有代表性。

统计对现象的研究要求具有总体性，也是基于满足统计研究的目的来考虑的。但强调总

体性的要求，并不排斥统计对社会经济个体现象观察的重视。事实上，统计对总体事物的研究是从对个体的观察开始的。例如在人口统计中，如果没有对一个自然人各方面情况的观察和记录，就得不到对人口总体的性别比例、民族分布、出生率、平均寿命等方面的数量认识。

3. 具体性

统计学研究具体地点、时间、条件下的社会经济现象的数量方面，这一特点是统计学与数学的根本区别，数学所研究的量是抽象的量，而统计学所研究的量是经济现象具体的量。

这是由于社会经济现象中的事物都是具体的，都是在一定的地点、时间、条件下发生的，所以其量的表现就必然带有特定场合和特定历史的痕迹，离开具体的地点、时间和条件，是无法说明社会经济现象的本质及其运行规律的。

但是，统计学毕竟是研究量的科学，虽然把统计学与数学区别开来，但这并不是说统计不需要利用数学研究方法。事实上，在实际统计工作中，研究具体的经济现象时，更要注意广泛利用各种数学分析方法进行研究。

4. 社会性

统计学通过社会经济现象总体数量的调查研究来认识人类社会活动的条件、过程和结果，反映物质资料的占有关系、分配关系、交换关系及其他社会关系。统计学研究的社会经济现象与各种利益关系是密切联系的。其定量研究是以定性分析为提前的，而定性分析使统计学在客观上有了社会关系的内涵。所以，统计学在研究社会经济现象时，必须注意正确处理好这些涉及人与人之间关系的社会矛盾。

5. 广泛性

统计学研究的数量方面非常广泛，包括整个社会，它既研究生产关系，也研究生产力及生产关系和生产力之间的关系；它既研究经济基础，也研究上层建筑及经济基础和上层建筑之间的关系。此外，还研究生产、流通、分配、消费等社会再生产的全过程及社会、政治、经济、军事、法律、文化、教育等全部社会现象的数量方面。

四、统计和统计学的产生与发展

（一）统计的产生与发展

统计作为一种社会实践活动，已有四五千年的历史，早在原始社会末期，人类就有了计数的概念和原始的计量方法，这可以说是统计的萌芽。当时统治阶级为了征税、征兵、服劳役的目的，需要对人口、土地、粮食和牲畜的数量进行计数，因此统计是适应社会发展和国家管理的需要而产生的。

我国早在公元前 21 世纪的夏朝就有了人口与土地数字的记载，当时全国分为九州，人口 1355 万人。而在公元前 27 世纪，埃及为了建造金字塔和大型农业灌溉系统，曾进行过全国人口和财产调查。公元前 6 世纪，罗马帝国规定每 5 年进行一次人口、土地、牲畜和家奴

的调查，并以财产总额作为划分贫富等级和征丁课税的依据。

进入封建社会以后，随着生产力的发展，统计的范围逐渐由人口、土地发展到社会经济生活的各个方面。由于自给自足的自然经济占主导地位，经济落后，长期的封建生产关系阻碍了社会生产力的提高，也阻碍了统计实践的发展。

统计实践活动的进一步发展始于资本主义社会。在 17～18 世纪资本主义上升时期，由于生产发展的需要，包括：人口、畜业、工业、农业、海关、外贸、物价等方面的统计，先后都得到了极大的发展。为了适应生产的发展和提升统计实践活动的需要，一门新兴的统计科学便应运而生了。

（二）统计学的产生与发展

17 世纪中叶，英国著名学者威廉·配第《政治算术》一书的问世，标志着古典政治经济学的诞生，也标志着统计学的诞生。从统计学诞生开始，许多人从不同的角度去认识和研究统计理论，逐渐形成不同的统计学派，它们同时共存、互相影响、互相争论，统计学就是在这种争论中逐步得到完善、充实和发展的。

在统计学的发展史上，主要的学派有政治算术学派、国势学派、数理统计学派和社会统计学派。

1. 政治算术学派

政治算术学派产生于 17 世纪中叶的英国，政治算术学派的代表人物是英国学者威廉·配第（William Petty，1623—1687）和约翰·格朗特（J. Graunt，1620—1674）。威廉·配第在其代表作《政治算术》中，运用数字、重量、尺度和数量对比分析的方法，为统计学的产生奠定了坚实的基础。正是在这个意义上，马克思称威廉·配第是“政治经济学之父，在某种程度上也可以说是统计学的创始人”。

英国人约翰·格朗特于 1662 年出版了《关于死亡表的自然观察与政治观察》。他根据伦敦市发表的人口自然变动公报，通过大量观察的方法，对人口的出生率和死亡率作了许多分类、计算和研究，发现了人口与社会现象中重要的数量规律性。在研究中，格朗特不但探索了人口变化和发展的一些数量规律，还对伦敦市总人口数量做出了较科学的估计。

政治算术学派一直未采用“统计学”这一科学名称，被后人认为是无统计学之名而有统计学之实的统计学正统。

2. 国势学派

国势学派产生于 17 世纪的德国，国势学派的代表人物是德国学者阿亨瓦尔（G．Achenwall，1714—1722）和康令（H．Conring，1606—1681）。国势学派所做的工作是对国家的重要事项进行记录，因此该学派又被称为记述学派，代表著作是《欧洲各国国势学概论》。阿亨瓦尔在大学中开设了一门新课程叫作“国势学”，因为在德文中“国势”与“统计”相通，后正式命名为“统计学”。从 1660 年开始，康令在西尔姆斯特大学以“国势学”

为题讲述一门课程，内容是各个国家的显著事项，方法则是文字叙述，目的是满足政治家所必需的知识。



小贴士

阿亨瓦尔

阿亨瓦尔是国势学的主要继承人和最有名的代表人物，一生在大学任教。他在 1749 年出版的《近代欧洲各国国势学论》中首次使用“统计学”（Statistik）这一名称代替国势学。阿亨瓦尔认为统计学是关于各国基本制度的学问，其研究对象是一个国家显著事项的整体。这里的“国家显著事项”是指一个国家的领土、人口、财政、军事、政治和法律制度等，用这些来说明和比较国家的形势，因此称为国势学。

国势学派搜集了各国大量的实际资料，分门别类，记述有关国情国力的系统知识，包括：土地、人口、政治、军事、财政、货币、科学、艺术和宗教等。但是始终没有把数量对比分析作为“统计学”的基本特征，被后人称为是有统计学之名而无统计学之实的统计学。

3. 数理统计学派

数理统计学派产生于 19 世纪中叶，创始人是比利时的天文学家、数学家和统计学家阿道夫·凯特勒（A.Quetelet，1796—1874），其著作有《统计学的研究》、《关于概率论的书信》等。凯特勒根据大数定律的原理提出了大量观察法，利用统计观察资料计算和研究社会现象和自然现象的数量规律性，并用于预测未来的情况。他创立大数法则，认为统计学就是数理统计学。

国际统计学界称凯特勒为“近代统计学之父”，由于这一学派主要在英美等国发展起来，故又称英美数理统计学派。数理统计学派最先将概率论应用于人口、人体测量和犯罪等问题的研究，完成了统计学和概率论的结合，发现了大量现象的统计规律并开创性地应用了许多统计方法，使统计方法产生了重大飞跃。从此，统计学开始进入更为丰富的新发展阶段。

4. 社会统计学派

社会统计学派于公元 19 世纪后半叶兴起于德国，主要代表人是恩格尔（C. L. E. Engel，1821—1896）和梅尔（G. V. Magr，1841—1925）。社会统计学派发展了政治算术学派的观点，他们主张统计学是研究社会现象的科学，研究方法为大量观察法。社会统计学派在理论上比政治算术学派更加完善，对国际统计学界影响较大，流传较广。

从 20 世纪 50 年代以来，统计理论、方法和应用进入了一个全面发展的新阶段。一方面，统计学受计算机科学、信息论、人工智能等现代科学技术的影响，新的研究领域层出不穷，如现代时间序列分析、非参数统计、线性统计模型等；另一方面，统计方法的应用领域不断扩展，几乎所有的科学研究都离不开统计方法。

第二节　统计学的研究方法和统计工作过程

一、统计学的研究方法

每门学科都有其特定的研究对象，不同的研究对象需要用不同的方法去研究。统计学是一门研究现象数量关系的方法论科学，长期以来，人们根据研究任务的要求，总结了一系列统计研究方法，如大量观察法、统计分组法、综合指标法、抽样推断法等，这里所讨论的是统计学研究中使用的基本方法。

1. 大量观察法

大量观察法是统计学中的特有方法。它是指统计在研究社会经济现象数量方面时，必须对总体现象中的全部或足够多的个体进行观察，以达到对现象总体数量特征及其规律性的认识。

大量复杂的社会经济现象是在诸多因素综合作用下形成的，个别单位往往受各种偶然因素的影响，使各单位的特征和数量表现有很大差别，所以不能任意抽取个别或少数单位进行观察。普查、抽样调查、统计报表调查、重点调查等则是大量观察的组织形式。

2. 统计分组法

统计分组法是根据统计研究问题的目的，选择不同的标准对总体进行不同的分组，以反映总体的构成和现象之间的依存关系。例如，要研究我国国有企业的有关情况，选择“企业规模”为标准进行分组，结果可以反映出国有企业中大、中、小型企业的数量和比例。

3. 综合指标法

综合指标法就是根据大量观察法获得的资料，计算、运用各种综合指标，以反映总体一般数量特征的统计分析法。通常使用的综合指标有总量指标、相对指标、平均指标、变异指标等。通过综合指标的计算可以显示出现象在具体时间、地点条件下的总量规模、相对水平、集中趋势、变异程度，并进一步从动态上研究现象的发展趋势和变化规律。

4. 抽样推断法

抽样推断法是指按照随机原则从总体中选择一部分单位进行调查，并根据结果对总体的数量特征做出有一定正确性和一定把握性估计的统计方法。这种方法主要用于难以进行全面调查的场合和不宜进行全面调查的场合。

除此以外，统计学研究中使用的其他方法还很多，如时间数列分析法、指数分析法和相关分析法等。在实际应用中，要根据具体情况，按照需要与可能，分别采用不同的研究方法，同时善于把多种研究方法结合应用，相互补充。

二、统计工作过程

统计工作是通过对社会经济现象进行调查研究，认识其本质和规律性的一种认识过程。统计认识过程和其他认识过程一样，是一个由感性认识到理性认识、不断发展与深化的过程。一项完整的统计工作可以分为四个阶段，即：统计设计、统计调查、统计整理和统计分析。

1. 统计设计

统计设计是根据统计研究对象的特点和研究的目的、任务，对统计工作的各个方面和各个环节进行通盘考虑和安排，确定调查对象的范围，明确分析对象的统计指标、指标体系和分组方法。统计设计既是统计认识过程的第一阶段，也是定性认识的阶段。统计设计的结果表现为各种设计方案，如调查方案的设计、整理方案的设计、统计分析方案的设计等。

统计设计的主要内容有：统计指标和指标体系的设计、统计分类和统计分组的设计、统计表的设计、统计资料搜集方法的设计、统计工作各个部门和各个阶段的协调与联系、统计力量的组织与安排。

2. 统计调查

统计调查是根据统计设计的要求，采用各种调查组织形式和调查方法，有组织、有目的地对所研究的对象进行观察、登记，准确、及时、系统、完整地搜集统计原始资料的过程。统计调查既是统计认识过程的第二个阶段，又是定量认识的阶段。统计调查的结果表现为各种调查表或登记表。

统计调查阶段所搜集的资料是否客观、全面、系统、及时，直接影响到统计整理结果的好坏，关系到统计分析结论的正确性，决定着整个统计工作的质量。所以，统计调查是整个统计工作的基础。

3. 统计整理

统计整理是指根据统计研究的目的，将统计调查得到的资料进行科学的分类和汇总，使其条理化、系统化的工作过程。通过统计调查搜集来的只是各个调查单位的分散的原始资料。这些大量的、分散的原始资料只能表明各个调查单位的具体情况，并不能说明总体的数量特征。为了了解现象的一般情况，显示总体的数量特征，接下来就是将按规定的统计指标体系调查来的统计资料进行整理。

统计整理是统计认识过程的第三阶段，是一种定量认识活动。统计整理既是统计调查的继续，也是统计分析的开始。统计整理的结果表现为各类统计图和统计表等。

4. 统计分析

统计分析是在统计整理的基础上，根据研究目的和任务，利用科学的统计分析方法，对统计研究对象的数量方面进行计算、分析的工作过程。统计分析的目的是揭示统计研究对象的状况、特点、问题和规律性，统计分析是统计认识的定性阶段。统计分析的结果表现为各类统计报告。

可见，统计工作的过程是经过统计设计到统计调查和统计整理，最后，通过统计分析而达到对事物本质和规律性的认识。这种质—量—质的认识过程，是统计工作的完整过程；缺少哪个环节，统计工作都会出现偏差。

需要注意的是，统计工作过程的四个阶段的划分在很大程度上只是理论上的，而在实践中，统计工作过程是很难这样清晰分开的。

第三节　统计学中的基本概念

统计学是一门方法论的学科，它和其他学科一样，在论述其理论和方法的过程中，经常要使用一些专门的术语和概念。统计学的基本概念包括：统计总体和总体单位、指标和标志、变异和变量等。

一、统计总体和总体单位

（一）统计总体

凡是客观存在的、在同一性质基础上结合起来的许多个别事物的整体，就是统计总体，简称总体。例如，要研究全国的运输能力情况，全国铁路、公路、船舶、航空及管道等所有具有运输能力的企业构成一个总体；又如，在全国的工业普查中，所有的工业企业构成一个总体。

总体具有以下三方面的特征。

1. 同质性

总体具有同质性，即构成总体的所有总体单位至少在一个方面具有相同的性质，必须具有某一方面的共性，这是构成总体的基础。例如，要调查我国工业企业的生产情况，全国的工业企业就构成一个总体，尽管这些企业规模大小、组织形式、产品属性和产量等各不相同，但它们都有一个共同的经济职能，即都是从事工业生产活动的。

同质性是相对的，它是根据研究目的而定的，目的不同，确定的总体就不同，同质性的意义也就改变了。例如，研究企业职工的工资水平，所有企业的职工构成了总体，凡属企业职工都是同质的。

2. 大量性

总体的大量性是指总体是由许多单位所组成的，构成总体的总体单位数量足够多，不能只是个别或少数个体。因为统计研究的目的是揭示现象的规律，个别单位的现象有很大的偶然性，而大量现象的总体则相对稳定，表现出共同性的倾向，这是现象的必然性。

总体根据所包含的单位数量，可以分为无限总体和有限总体。无限总体是指总体所包含

的单位数很多，乃至无限多。例如，研究海洋中鱼类的生长情况，海洋中的全部鱼类难以计数，就是无限总体。有限总体是指总体中所包含的单位数量是有限的，是可以计数的。例如，某一时点上的人口数、工业企业的个数、校园里的学生人数等所组成的总体就都是有限总体。

3. 差异性

总体的差异性是客观存在的，即总体的各个单位之间，除了必须在某一方面有共性之外，在其他方面还存在着差异，这些差异是统计研究的基础。例如，在运输企业总体中，每个企业除了具有运输能力的共同性质外，其他方面如职工人数、所有制性质、固定资产、销售收入、利润等都各不相同。

差异是客观的，没有差异的事物是不存在的。没有差异就没有统计，这是因为如果总体单位之间不存在差异，我们只需要了解一个总体单位的资料就可以推断总体情况了。这种差异是统计研究的前提，统计研究就是要在个别事物的差异中寻找共性，揭示其活动规律。差异性和同质性之间相互联系、相互补充，是辩证统一的关系。

必须同时具备上述三个特征，才能够形成总体，由此可以进行一系列的统计计算和统计分析。

（二）总体单位

构成统计总体的个别事物称为总体单位。总体单位是构成总体的基础，要了解总体的数量特征，就必须从总体单位一个个登记调查开始。全国工业普查中的各个工业企业都是总体单位。如果把“全国总人口”作为一个总体，那么，每一个具有中华人民共和国国籍的公民都是总体单位。

（三）总体和总体单位的相互关系

总体和总体单位的概念是相对而言的，不是固定不变的，随着研究目的不同，总体和总体单位的确认也会有所不同。例如，在工业企业这个统计总体的例子里，每个工业企业都是一个总体单位，但是，当研究一个工业企业中职工工资情况时，则该企业的所有职工就成为统计总体，各个职工是总体单位。

二、统计指标与标志

（一）统计指标

1. 统计指标的含义

统计指标简称指标，从统计理论上讲，统计指标是反映总体现象数量特征的概念。例如国内生产总值、人均工资、人口数、劳动生产率等。这种含义的统计指标包含三个要素，即指标名称、计量单位和计算方法。

在实际统计工作中，使用的是另一种含义的统计指标，这种含义的统计指标是指反映总体现象数量特征的概念及其数值，如 2016 年上半年京津冀地区生产总值合计 34 382.4 亿元，

这时统计指标除包含上述三个要素外，还包含了时间限制、空间限制和指标数值三个要素。

不同含义统计指标在不同的场合下使用，在进行统计设计时，只能设计指标的名称、内容、口径、计量单位和计算方法，这是不包括数值的统计指标；然后经过搜集资料、汇总整理、加工计算，可以得到统计指标的具体数值，用来反映总体现象的实际数量状况及其发展变化情况。

2. 统计指标的种类

从不同角度，可以将统计指标分为以下几个种类。

（1）按统计指标所说明的总体现象内容分类

按统计指标所说明的总体现象内容的不同，可以分为数量指标和质量指标。

数量指标又称为总量指标，是反映总体绝对数量多少的指标，用绝对数的形式来表示，并有计量单位。数量指标的数值随总体外延范围的大小而增减，如全国的人口数大于某一省的人口数。数量指标可表现为总体总量，即一个总体中单位的数目，如企业数也可表现为标志总量，即总体各单位某一标志值的总和，如某企业职工的工资总额。

质量指标是说明总体内部数量关系和总体单位水平的指标，它通常以相对数和平均数的形式来表示，如劳动生产率、平均工资等。质量指标的数值不随总体范围的变化而变化。例如，2015 年全国城镇私营单位就业人员年平均工资为 39 589 元；分四大区域看，城镇私营单位就业人员年平均工资由高到低依次是东部、西部、中部和东北，分别为 43 439 元、36 478 元、32 773 元和 32 176 元。也就是说，平均工资是一个质量指标，它与总体规模的大小没有直接关系。

（2）按统计指标表现形式分类

按统计指标表现形式不同，可以分为总量指标、相对指标和平均指标。

总量指标是反映总体规模的统计指标，表明总体现象发展的结果。例如人口总数、国内生产总值、工资总额等，都是总量指标。

相对指标反映现象之间的对比关系，如发展速度、计划完成相对数等。平均指标反映现象总体内部各单位的一般水平，如平均工资、平均年龄、粮食平均亩产量等。

（3）按指标所反映现象的时间状况分类

按指标所反映现象的时间状况不同，可以分为静态指标和动态指标。

静态指标反映既定时间上现象的规模、水平、数量关系等，如一般平均数、绝对数和许多相对数等。动态指标反映现象在不同时间内发展变化的情况，如序时平均数、发展速度、增长速度等。

（4）按统计指标的作用分类

按指标所反映现象的时间状况不同，可以分为静态指标和动态指标。

描述指标是用来反映现象基本情况的指标，如医生总数、学生总数、外汇储备总额等。

评价指标是用来对客观事物活动的结果进行评估和考核的指标，如对工业企业经营活动效益进行评价的产品销售率、流动资金周转率、劳动生产率等指标。

预警指标是对现象宏观运行进行监测，并据此对可能出现的总体失衡、结构性矛盾、突

发异常情况作出预报的指标，如通货膨胀率、失业率、人口增长率等，这类指标涉及面广，敏感性强，对国民经济的发展和社会稳定具有重要作用。

此外，统计指标还可以从其他角度进行分类，如按指标的时间标准不同可以为时点指标和时期指标；按指标的计量单位不同分为实物指标、价值指标和劳动量指标等。

3. 统计指标体系

（1）统计指标体系的含义

由于一个统计指标只能反映社会经济现象的某一个侧面的特征，而社会经济现象有许多侧面和特征，它们相互依存制约。为了全面、深入地认识社会经济现象，就必须采用多个具有相互联系的统计指标来综合反映社会经济现象的整体特征。

将由若干个相互联系的统计指标组成的整体称为统计指标体系，如用劳动生产率、资金利税率、工业增加值等一系列指标来研究工业企业的全面情况，这些指标就组成了工业企业的统计指标体系。

指标体系的设置不但是客观现象的反映，也是人们对客观认识的结果。随着客观形势的发展变化及实践经验和理论研究的积累，指标体系也将不断改进更新，渐臻完善。

（2）统计指标体系的种类

①统计指标体系按其所反映的内容不同，可以分为基本统计指标体系和专题统计指标体系。基本统计指标体系是经常系统地反映生产、分配、交换、消费的经济指标体系，反映文化、教育、卫生的社会指标体系，反映科学技术人员、投入和成果的科技指标体系等。专题指标体系是为对某一专门问题进行统计研究而设计的统计指标体系，如反映农村贫困状况的指标体系、反映小康的指标体系、地区可持续发展指标体系等。

②统计指标体系按其实施范围不同，可以分为国家统计指标体系、行业（或部门）统计指标体系、地方统计指标体系、基层单位的统计指标体系。

（二）标志

标志是说明总体单位所具有的属性或特征的，包括标志名称和标志表现两部分。每个总体单位都有许多属性和特征，如企业中每一名职工作为总体单位考察时，都具有性别、民族、文化程度、年龄、工资等属性和特征，这些都属于标志名称，而这些标志名称各有一定的具体表现，如性别表现为“男”或“女”，这些就是标志表现。

标志有品质标志与数量标志之分。品质标志表示事物的品质属性特征，是不能用数值表示的，如性别、民族等；数量标志表现事物的数量特征，是可以用数值表示的，如年龄、工资等。各种标志都有自己的标志表现，数量标志表现为具体的数值，如在人口研究中某人年龄 30 岁、体重 56 公斤等，这里的“30”和“56”都是标志值；品质标志表现为对特征加以描述的文字，如民族“汉”等。

在一个总体中，每个总体单位都具有不变标志与可变标志。不变标志是指对所有总体单位都有完全相同标志表现的标志；正因为具有这个不变标志，才使它们集合在一起构成同质总体。可变标志是指在总体单位之间具有不同标志表现的标志。

例如，对某地区所有工业企业这个总体来说，其不变标志是“某个地区”、“工业”，这

两个标志对总体各单位进行了具体的界定，构成企业的同质性；而每个企业的职工人数、产量、产值等都可能不同，是可变标志，它们构成总体单位的变异性。总体的同质性和总体单位的变异性是进行统计核算的条件。

统计总体、总体单位、指标、标志的关系如图 1-1 所示。

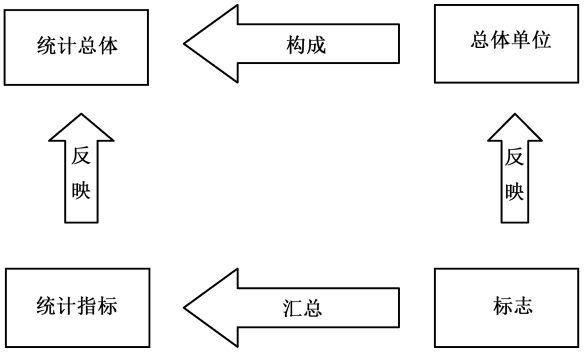


图 1-1　统计总体、总体单位、指标、标志的关系

（三）统计指标与标志的关系

统计指标与标志既有明显的区别，又有密切的联系。

两者的区别主要表现在以下方面。

①指标是说明总体特征的：而标志是说明总体单位特征的。

②标志可以分为不能用数值表示的品质标志与能用数值表示的数量标志，而指标都是用数值表示的，没有不能用数值表示的指标。

③标志中的数量标志不一定经过汇总，可以直接取得；而指标是由数量标志汇总得来的。

④标志一般不具备时间、地点等条件；而作为一个完整的统计指标，一定要有时间、地点、范围。

两者的联系主要表现在以下方面。

①许多指标的数值是从总体单位的数量标志值汇总而来的。

②指标与数量标志之间存在着变换关系。由于研究目的不同，总体和总体单位也会不同，有的指标可能会变成标志，有的数量标志也可能变成指标。例如，当北京市 18 个区县人口数构成总体时，则密云县的人口数是标志；而当密云县所有乡镇人口数构成总体时，密云县的人口数就是指标，因为它反映的是总体的数量特征。

三、变异与变量

（一）变异

每个总体单位都具有不同的特征，以区别于另一总体单位。标志在不同总体单位之间不断变化，由一种状况变为另一种状况，这种变化就是变异。例如，人的性别标志表现为男、

女；年龄标志表现为 20 岁、30 岁等。变异是普遍存在的，这是统计的前提条件。

（二）变量和变量值

在统计研究中，变量是一个非常重要的概念。变量是现象发展变化的数量化概念，或者说是现象本身所固有的、随条件变化而变化的量。我们前面讲过的所有统计指标和数量标志都是变量。变量的具体数值表现就是变量值。

变量根据不同的划分标准可以有以下几种分类。

1. 按数值的连续性不同分类

变量按数值的连续性不同，可以分为连续变量与离散变量。连续变量的数值是连续不断的,相邻的两个数值之间可作无限分割，即可取无限个数值。例如人的身高、体重、劳动生产率等。连续变量的数值要用测量或计算的方法取得。

离散变量的数值是以整数位断开的，相邻的两个变量值之间不可能再分割出新的变量值，这样的变量称为离散变量，如职工人数、工业企业数、产品件数等。离散变量的数值只能用计数的方法取得。

2. 按事物的性质不同分类

变量按事物的性质不同，可以分为确定性变量和随机性变量。确定性变量是指变量受某些确定性因素影响，现象的量会沿着某一方向持续变化。例如，由于科学技术的不断提高和医疗卫生条件的不断改善，人类的死亡率在逐步降低，人类的平均寿命在不断延长，因此从长期来看，人的平均寿命和死亡率都是确定性变量。

随机性变量是指变量的变动受许多因素变动的影响，变量值的大小没有明确的方向，结果带有偶然性。例如，按随机原则从总体中选取容量一定的样本，每一次都会得到不同的结果，因此，样本是个随机变量。

3. 按事物的类型不同分类

变量按事物的类型不同可以分为分类变量、顺序变量和数值变量。

（1）分类变量

分类变量是说明事物类别的一个名称。例如，“性别”就是一个分类变量，其变量值表现为“男”或“女”;“宗教信仰”也是一个分类变量，其变量值表现为“佛教”、“天主教”、“伊斯兰教”等。分类变量的变量值用文字表示，变量值顺序可以改变。

（2）顺序变量

顺序变量（次序变量）是说明事物有序类别的一个变量，顺序变量的值是有序的。如“受教育程度”是一个顺序变量，其变量值可以表现为“小学”、“初中”、“高中”、“大学”等；一个人对某种事物的看法也可以是一个顺序变量，其变量值可以表现为“满意”、“较满意”、“不满意”等。

（3）数值变量

数值变量是说明事物数量特征的一个名称。例如，“家庭子女数”、“城镇登记失业率”、“年龄”、“时间”等都是数值变量。数值变量有测量单位，可以说出一个值比另一个值大多少

或小多少。

我们可以把数值变量转换成分类变量或顺序变量，但反过来不行。例如，收入水平可以转换成贫困、温饱、小康和富裕的顺序变量。但贫困、温饱、小康和富裕不能变成数值型的收入水平。分类变量、顺序变量和数值变量的特点如表 1-1 所示。

表 1-1　分类变量、顺序变量和数值变量的特点

分 类 变 量	顺 序 变 量	数 值 变 量
顺序可以改变	顺序不可以改变	可以进行各种数学运算
不可以比较大小	可比较大小	
不可以进行数学运算	不可以进行数学运算	
性别：男、女	考试成绩：优、良、中、差	年龄：11 岁、12 岁、13 岁



本章小结

统计一词有统计工作、统计资料和统计学三种含义，三者之间存在着密切的联系。统计学研究大量社会经济现象的数量方面，具有数量性、总体性、具体性、社会性、广泛性五大特点，统计学研究的基本方法有大量观察法、统计分组法、综合指标法、抽样推断法等。

统计工作过程一般包括统计设计、统计调查、统计整理、统计分析四个阶段。统计有信息、咨询和监督三大职能。统计总体是客观存在的，总体单位则是构成总体的基础。

统计指标反映现象总体特征的数量概念和具体数值。标志是说明总体单位的属性特征的，指标与标志既有联系又有区别。

变异是总体单位间同一标志的差别，所有统计指标和所有数量标志都是变量。变量的具体数值表现就是变量值。变量有离散变量和连续变量等分类。



复习思考题

1．统计一词有几种含义？它们之间有什么关系？

2．统计研究有哪些方法？

3．统计工作的过程包含哪几个阶段？

4．统计指标和标志有什么联系和区别？

5．什么是变异和变量？举例说明变量的种类？

6．怎样理解统计研究的对象及其特点？



实训基础训练

一、单项选择题

1．一个统计总体（　　）。

A．只能有一个标志　　B．只能有一个指标

C．可以有多个标志　　D．可以有多个指标

2．下列标志中，数量标志是（　　）。

A．社会阶层　　B．性别　　C．民族　　D．汽车产量

3．下列总体中，属于无限总体的是（　　）。

A．全国人口总数　　B．水塘中的鱼

C．城市年流动人口数　　D．工业中连续大量生产的产品产量

4．某小组学生语文考试成绩分别为 90 分、98 分、85 分和 95 分。这四个数字是（　　）。

A．标志　　B．指标　　C．标志值　　D．变量

5．某地区共有 30 所学校，考察这些学校学生的学习总体情况，则统计总体是（　　）。

A．30 所学校的全体学生　　B．30 所学校

C．30 所学校全体学生的成绩　　D．30 所学校学生的平均成绩

6．要了解某企业职工文化程度，其总体单位是（　　）。

A．该企业全部职工　　B．该企业每一名职工

C．该企业每一名职工的文化程度　　D．该企业全部职工的平均文化程度

7．工业企业的设备台数、产品产值（　　）。

A．都是连续变量　　B．都是离散变量

C．前者是连续变量，后者是离散变量　　D．前者是离散变量，后者是连续变量

8．对某公司 200 名职工的收入状况进行调查，全体职工的平均年龄是（　　）。

A．数量标志　　B．数量指标　　C．品质标志　　D．质量指标

9．指标是说明总体特征的，标志是说明总体单位特征的，所以（　　）。

A．标志和指标之间的关系是固定不变的

B．标志和指标之间的关系是可以变化的

C．只有标志可以用数值表示的

D．只有指标可以用数值表示

10．数量指标的表现形式是（　　）。

A．绝对数　　B．相对数　　C．平均数　　D．小数

二、判断题

1．统计学是观察研究大量现象数量方面的科学。（ ）
2．在全国工业普查中，全国工业企业数是统计总体，每个工业企业是总体单位。（ ）
3．统计学的研究对象是社会经济现象的各个方面。（ ）
4．总体和总体单位是永远不变的。（ ）
5．对某市常住人口情况进行调查，则常住人口数是连续型变量。（ ）
6．社会经济现象中的统计总体大都属于有限总体。（ ）
7．数量指标是数值形式，质量指标不是数值形式。（ ）
8．对某校学生的身高进行调查，则学生的身高是离散型变量。（ ）
9．质量指标通常以相对数和绝对数的形式表现。（ ）
10．职工的年龄是品质标志。（ ）



课外延伸阅读

国家统计主要指标解释
——人民生活

一、城镇住户

城镇家庭人口：指居住在一起，经济上合在一起共同生活的家庭成员。凡计算为家庭人口的成员其全部收支都包括在本家庭中。

城镇就业面：指就业人口占家庭人口的百分比。

城镇就业者负担人数：指家庭人口与就业人口之比。

城镇家庭总收入：指家庭成员得到的工资性收入、经营净收入、财产性收入、转移性收入之和，不包括出售财物收入和借贷收入。

城镇居民家庭可支配收入：指家庭成员得到可用于最终消费支出和其他非义务性支出及储蓄的总和，即居民家庭可以用来自由支配的收入。它是家庭总收入扣除交纳的个人所得税、个人缴纳的社会保障支出及记账补贴后的收入。计算公式为：

城镇居民家庭可支配收入=家庭总收入-缴纳个人所得税-
个人交纳的社会保障支出-记账补贴

城镇家庭总支出：指家庭除借贷支出以外的全部实际支出。包括现金消费支出、财产性支出、转移性支出、社会保障支出、购房与建房支出。

城镇家庭现金消费支出：指家庭用于日常生活的全部现金支出，包括食品、衣着、居住、家庭设备及用品、交通通信、文教娱乐、医疗保健、其他等八大类支出。

城镇家庭服务性消费支出：指家庭用于支付社会提供的各种文化和生活方面的非商品性

服务费用。

城镇家庭收入分组方法：是将所有调查户按户人均可支配收入由低到高排队，按 10%、10%、20%、20%、20%、10%、10%的比例依次分成：最低收入户、较低收入户、中等偏下收入户、中等收入户、中等偏上收入户、较高收入户、最高收入户七组。总体中最低 5%的户为困难户。

恩格尔系数：指食品支出在现金消费支出中所占的比例。

二、农村住户

农村住户：指农村常住户。农村常住户指长期（一年以上）居住在乡镇（不包括城关镇）行政管理区域内的住户，以及长期居住在城关镇所辖行政村范围内的农村住户。户口不在本地而在本地居住一年及以上的住户也包括在本地农村常住户范围内；有本地户口，但举家外出谋生一年以上的住户，无论是否保留承包耕地，都不包括在本地农村住户范围内。

常住人口：指全年经常在家或在家居住 6 个月以上，而且经济和生活与本户连成一体的人口。外出从业人员在外居住时间虽然在 6 个月以上，但收入主要带回家中，经济与本户连为一体，仍视为家庭常住人口；在家居住，生活和本户连成一体的国家职工、退休人员也为家庭常住人口。但是现役军人、中专及以上（走读生除外）的在校学生及常年在外（不包括探亲、看病等）且已有稳定的职业与居住场所的外出从业人员不算家庭常住人口。家庭常住人口主要作为计算农村住户平均每人收入、消费和积累水平及分析家庭人口状况的依据。

整、半劳动力：整劳动力指男子 18 周岁到 50 周岁，女子 18 周岁到 45 周岁；半劳动力指男子 16 周岁到 17 周岁，51 周岁到 60 周岁；女子 16 周岁到 17 周岁，46 周岁到 55 周岁，同时具有劳动能力的人。虽然在劳动年龄之内，但已丧失劳动能力的人，不应算为劳动力；超过劳动年龄，但能经常参加劳动，计入半劳动力数内。常住人口中的职工，若这些职工为劳动力，就包括在本户的整半劳动力中。

总收入：指调查期内农村住户和住户成员从各种来源渠道得到的收入总和。按收入的性质划分为工资性收入、家庭经营收入、财产性收入和转移性收入。

工资性收入：指农村住户成员受雇于单位或个人，靠出卖劳动而获得的收入。

家庭经营收入：指农村住户以家庭为生产经营单位进行生产筹划和管理而获得的收入。农村住户家庭经营活动按行业划分为农业、林业、牧业、渔业、工业、建筑业、交通运输业邮电业、批发和零售贸易餐饮业、社会服务业、文教卫生业和其他家庭经营。

财产性收入：指金融资产或有形非生产性资产的所有者向其他机构单位提供资金或将有形非生产性资产供其支配，作为回报而从中获得的收入。

转移性收入：指农村住户和住户成员无须付出任何对应物而获得的货物、服务、资金或资产所有权等，不包括无偿提供的用于固定资本形成的资金。一般情况下，指农村住户在二次分配中的所有收入。

现金收入：指农村住户和住户成员在调查期内得到以现金形态表现的收入。按来源分成工资性收入、家庭经营现金收入、财产性收入、转移性收入。

农村居民家庭纯收入：指农村住户当年从各个来源得到的总收入相应地扣除所发生的费用后的收入总和。计算方法：

农村居民家庭纯收入=总收入-家庭经营费用支出-税费支出-
生产性固定资产折旧-赠送农村内部亲友

纯收入主要用于再生产投入和当年生活消费支出，也可用于储蓄和各种非义务性支出。“农民人均纯收入”是按人口平均的纯收入水平，反映的是一个地区农村居民的平均收入水平。

总支出：指农村住户用于生产、生活和再分配的全部支出。包括家庭经营费用支出、购置生产性固定资产支出、税费支出、消费支出、财产性支出和转移性支出。

资料来源：中华人民共和国国家统计局网站
http://www.stats.gov.cn/tjsj/zbjs/201310/t20131029_449516.html



第2章 统计调查

学习目标要求

1. 了解统计调查的概念、任务、要求及基本原则，掌握统计调查方案的制定。
2. 掌握统计调查的种类、特点及适用条件。

重点难点提示

1. 统计调查的方法，调查方案涵盖的内容。
2. 统计报表、普查、典型调查、重点调查、抽样调查的含义以及适用性。

教学引导案例

第三次全国经济普查主要数据公报（第1号）

根据《国务院关于开展第三次全国经济普查的通知》（国发〔2012〕60号）要求，我国进行了第三次全国经济普查。这次普查的标准时点为2013年12月31日，普查时期资料为2013年年度资料。普查对象是在我国境内从事第二产业和第三产业的全部法人单位、产业活动单位和个体经营户。通过这次普查，摸清了我国第二产业和第三产业的发展规模及布局，摸清了我国产业组织、产业结构、产业技术的现状以及各生产要素的构成，查实了服务业、战略性新兴产业、小微企业和高技术产业（制造业）的发展状况。通过对31个地区的数据质量抽查，数据填报综合差错率为3.3‰，普查数据质量达到预期目标要求。

一、单位基本情况

2013年末，全国共有从事第二产业和第三产业活动的法人单位1085.7万个，比2008年末（2008年是第二次全国经济普查年份，下同）增加375.8万个，增长52.9%；产业活动单位1303.5万个，增加417.1万个，增长47.1%；有证照个体经营户3279.1万个，增加405.4万个，增长14.1%。

2013年末，在第二产业和第三产业法人单位中，位居前三位的行业是：批发和零售业281.1万个，占25.9%；制造业225.3万个，占20.7%；公共管理、社会保障和社会组织152万个，占14%。在有证照个体经营户中，位居前三位的行业是：批发和零售业1642.7万个，占50.1%；交通运输、仓储和邮政业878.6万个，占26.8%；住宿和餐饮业240.8万个，占7.3%。

2013 年末，全国共有第二产业和第三产业的企业法人单位 820.8 万个，比 2008 年末增加 324.9 万个，增长 65.5%。其中，内资企业占 97.5%，港、澳、台商投资企业占 1.2%，外商投资企业占 1.3%。内资企业中，国有企业占全部企业法人单位的 1.4%，私营企业占 68.3%。

二、从业人员

2013 年末，全国第二产业和第三产业法人单位从业人员 35 602.3 万人，比 2008 年末增加 8290.8 万人，增长 30.4%。有证照个体经营户从业人员 9013.4 万人，比 2008 年末增加 818 万人，增长 10%。

在法人单位从业人员中，位居前三位的行业是：制造业 12 515.1 万人，占 35.2%；建筑业 5320.6 万人，占 14.9%；批发和零售业 3315 万人，占 9.3%。在有证照个体经营户从业人员中，位居前三位的行业是：批发和零售业 4166.6 万人，占 46.2%；交通运输、仓储和邮政业 1674.5 万人，占 18.6%；住宿和餐饮业 1069.4 万人，占 11.9%。

三、企业资产总计

2013 年末，全国第二产业和第三产业企业资产总计 466.8 万亿元。其中，第二产业企业资产总计占全部企业资产总计的 26.1%，第三产业企业资产总计占 73.9%。

四、小微企业

2013 年末，全国共有第二产业和第三产业的小微企业法人单位 785 万个，占全部企业法人单位 95.6%。其中，位居前三位的行业是：工业 234.2 万个，占全部企业法人单位 28.5%；批发业 169.8 万个，占 20.7%；零售业 103.1 万个，占 12.6%。

小微企业从业人员 14 730.4 万人，占全部企业法人单位从业人员 50.4%。其中，位居前三位的行业是：工业 7403.6 万人，占全部企业法人单位从业人员 25.4%；建筑业 1675.4 万人，占 5.7%；批发业 1457.8 万人，占 5%。

小微企业法人单位资产总计 138.4 万亿元，占全部企业法人单位资产总计 29.6%。其中，位居前三位的行业是：工业 40.8 万亿元，占全部企业法人单位资产总计 8.7%；租赁和商务服务业 36.8 万亿元，占 7.9%；房地产开发经营 18.7 万亿元，占 4%。

五、战略性新兴产业

2013 年末，在第二产业和第三产业企业法人单位中，有战略性新兴产业活动的企业法人单位 16.6 万个，占全部企业法人单位 2%。其中，节能环保产业 7.1 万个，占全部企业法人单位 0.9%；新材料产业 4.7 万个，占 0.6%。

有战略性新兴产业活动的企业法人单位从业人员 2362.3 万人，占全部企业法人单位从业人员的 8.1%。其中，节能环保产业 1003.9 万人，占全部企业法人单位从业人员的 3.4%；新材料产业 707.9 万人，占 2.4%。

资料来源：中华人民共和国国家统计局网站

http://www.stats.gov.cn/tjsj/zxfb/201412/t20141216_653709.html

思考：

1. 为什么要开展统计调查？
2. 如何开展统计调查，需要做哪些准备？

第一节　统计调查概述

统计调查是统计整理和统计分析的基础。通过统计调查取得丰富的实际资料，是统计整理和统计分析的前提，是保证统计工作质量的基本环节。统计调查是搜集资料、获得感性认识的阶段，它是对现象总体认识的开始，只有搞好统计调查，才能保证统计工作达到对客观事物规律性的认识，并可以从中预测未来。

如果调查的资料不完整、不真实或不及时，即使是科学地整理、周密地分析，也不可能得到正确的判断，这将直接影响整个统计工作的成果。统计调查在很大程度上直接影响着统计工作任务完成的好坏，决定着整个统计工作质量的优劣。

一、统计调查的概念和基本原则

（一）统计调查的概念

统计数据是研究实际问题必不可少的重要资料。统计调查则是从统计信息源上采集基础数据的方法，是具体认识客观事物的起点。统计调查是按照预定的调查方案，运用科学的调查方法，有组织、有计划地及时向客观实际搜集统计资料的工作过程。

统计调查的基本任务是取得反映社会经济现象总体全部或部分单位以数字资料为主体的信息。与一般的社会调查比较：相同点是同属于调查研究活动，区别是统计调查要搜集大量的、以数字资料为主体的信息。

在统计实践中，我们搜集到的基础资料有两类：一类是初级资料，又称原始材料，是直接向调查对象搜集到的反映调查单位特征的未经任何加工整理的统计资料；另一类是次级资料，是指从统计台账、内部报表、财务报表、统计报表中搜集到的已经经过加工整理的统计资料。统计调查的主要资料来源是初级资料。

（二）统计调查的基本原则

为了保证调查资料的质量，统计调查应该遵循以下三个原则：要求统计调查做到准确、及时和全面，这是衡量统计工作质量的重要标志。

1. 准确性

统计调查所取得的资料必须如实反映客观实际。只有资料准确，才能正确地反映被研究现象的真实情况，对这些资料进行整理和分析，才能得出正确的结论。所以，应把调查资料的准确性当作统计工作的生命线来看待。

2. 及时性

及时性即时效性，是指在统计调查规定的时间以内完成统计调查工作并提供统计资料。

统计信息的时效性是统计工作的价值所在，失去时效性的资料犹如“雨后送伞”、“亡羊补牢”。因此，各级统计部门都应及时向各级领导提供所需资料，从时间上满足各层次对统计资料的需求，将及时性当作统计工作的重中之重。

3. 全面性

全面性就是对要调查的单位和项目，必须毫无遗漏地搜集资料，只有系统地从不同方面搜集反映现象发展过程、特征及问题的全面材料，才能对所研究现象总体的全貌加以如实反映；反之，如果资料不全，就不能全面反映调查对象的全貌，甚至会得出错误的结论。

二、统计调查的种类

由于社会经济现象的复杂性和统计研究任务的多样性，为了准确搜集原始资料，应该根据统计研究的目的和调查对象的特点采用不同的调查方式和方法。常用的统计调查方式主要有以下几类。

（一）按照调查对象分类

按调查对象范围的不同，统计调查可分为全面调查和非全面调查。

1. 全面调查

全面调查是对构成调查对象的所有单位进行逐一的、无一遗漏的调查，包括全面统计报表和普查。

全面调查的目的在于掌握所有调查单位的全面情况，如全国经济普查、人口普查、农业普查等。全面调查调查的项目较多，取得较细致、较深入的材料，但因需要耗用的人力、物力、财力较多，花费的时间较长，出现调查误差的可能性较大，因此在不影响统计研究目的实现的条件下，一般可以采用非全面调查的形式。

2. 非全面调查

非全面调查是对调查对象中的一部分单位进行调查，包括非全面统计报表、抽样调查、重点调查和典型调查。

由于非全面调查方式调查的单位少，如为了了解城市居民家庭收入水平，从 10 万户中随机抽选 5000 户进行抽样调查，这样可以节约大量的人力、物力、财力，花费的时间较短，提高了统计资料的时效性。

（二）按照调查登记时间分类

按调查登记的时间是否连续，统计调查分为经常性调查和一次性调查。

1. 经常性调查

经常性调查是指对研究对象的变化进行连续不断的登记，观察事物在一段时期内的发展变化。例如，工业企业总产值、产品产量、原材料消耗量等，在观察期内连续登记。经常性

调查所得的资料是现象在一段时间内的总量。

经常性调查是指调查周期在一年以内的调查，间隔超过一年的调查为一次性调查。这种划分和调查对象没有关系，不要把经常性调查误以为是全面调查，也不要误以为经常性调查就是调查时期现象而一次性调查就是调查时点现象。

2. 一次性调查

一次性调查是指间隔一段相当长的时间对研究对象某一时刻的资料进行登记。如果人口数、机器设备台数等资料短期内变化不大，就没有必要连续登记资料。不连续调查所得的资料体现现象在某一瞬间所具有的水平。



小贴士

一次性调查

一次性调查可以是定期进行的，也可以是不定期进行的。例如，2000 年 11 月 1 日零时进行了全国第五次人口普查，2010 年 11 月 1 日零时进行了全国第六次人口普查。今后，我国的普查将规范化、制度化，即每逢末尾数字为“0”的年份进行人口普查，每逢“5”的年份进行经济普查。

（三）按照调查组织方式分类

按调查组织的方式不同，统计调查可分为统计报表制度和专门调查。

1. 统计报表制度

统计报表制度是国家统计系统和专业部门为了定期取得系统、全面的统计资料而采用的一种搜集资料的方式，目的在于掌握经常变动的、对国民经济有重大意义的指标的统计资料。

2. 专门调查

专门调查是为了解和研究某种情况或问题而专门组织的统计调查，包括抽样调查、普查、重点调查和典型调查等几种调查方法。

（四）按照调查搜集资料的方法分类

按搜集资料的方法不同，统计调查分为直接调查法和间接调查法两大类。

1. 直接调查法

直接调查法主要包括以下方法。

（1）直接观察法

直接观察法是由调查人员在现场运用感觉器官或借助科学仪器对调查对象进行观察计量，以取得第一手调查资料的方法。例如，库存商品的盘点、车流量的计数调查等。

此种方法简便易行，调查取得的资料一般比较准确，但需要投入的人力多、时间长。

（2）询问法

询问法也称采访法，是由调查人员以调查表或有关材料为依据，逐项向被调查者询问有关情况并将答案记录下来，根据被调查者的回答取得调查资料的一种方法。例如，人口普查中的人口登记等。询问法又分为口头询问和被调查者自填两种。

此种调查方法取得资料的准确性取决于调查人员和被调查者的配合程度及被调查者的态度，优点是调查人员与被调查者可以直接交流，但需要的人力多，花费也较多。

（3）报告法

报告法是以各种原始和核算凭证为调查资料来源，依照统一的表格形式和要求，按照隶属关系，逐级向有关部门提供资料的方法。采用这种方法比较省时省力，还可以促使被调查者建立健全的原始资料，但容易出现虚报瞒报的现象。

（4）问卷法

问卷法是一种以问卷形式提问、由被调查者自愿回答，调查者将答案汇总而搜集资料的方法。问卷法是应用范围比较广泛的一种调查方法。在统计研究中，问卷法或单独使用，或与其他方法（如采访法、直接观察法）同时使用。它既适用于理论性问题的调查，也适用于应用性问题的调查；既适用于小规模调查，也适用于大规模调查。

问卷法的关键在于精心设计问卷，为保证问卷调查的效果，提问要简明扼要，答案要标准化，易于选择和汇总。这种方法省时省力，但如果回收率较低或反馈的答案质量不高，就会影响数据的质量及结论的可靠性，因此，在实施上要尽量设法防止回收率较低或反馈答案质量不高的问题。

2. 间接调查法

间接调查法是指调查人员收集二手资料的方法。调查人员可以查阅文献、报刊、统计报告和统计年鉴等获得资料，也可以通过广播、电话、电视、网络等通信手段获得信息。目前，网络搜索是一种快捷、便利、高效的获取资料的形式。

第二节　统计调查方案

统计调查工作，尤其是大规模的统计调查工作，调查的内容多、范围广，需要大量人员参与才能完成任务。为了保证统计调查工作顺利进行并能达到预期的目标，在统计调查之前，要制定一个科学合理的统计调查方案。

统计调查方案是统计调查前所制订的实施计划，是全部调查过程的指导性文件。是调查工作有计划、有组织、有系统地进行的保证。不同调查任务的调查方案在具体内容和形式上会有一定的差别，但包括的主要内容大体是一致的。统计调查方案应确定的内容有：调查的目的与任务、调查对象与调查单位、调查项目与调查表、调查时间和调查时限、调查的组织

实施计划。

一、调查目的与任务

确定调查目的与任务是任何一项统计调查方案首先要确定的问题。它回答的是：为什么调查？要解决什么样的问题？调查具有什么样的社会经济意义？如果目的不明确、任务不清楚，就无法确定向谁做调查、调查什么，以及用什么方法进行调查。其结果不仅浪费人力、物力、财力，同时，由于搜集的资料不适宜，还将导致整个统计工作的延误。

调查目的是指调查所要达到的具体目标。调查目的的写作应简明扼要。不同的调查目的需要不同的调查资料，不同的调查资料又有不同的搜集方法。调查目的明确了，搜集资料的范围和方法也就确定下来了。

统计调查的目的决定了调查任务，统计调查的基本任务就是按照所确定的指标体系，通过具体的调查，取得反映社会经济现象总体全部或部分单位以数字资料为主体的信息。



小贴士

第三次全国经济普查的主要目的

2014 年 12 月 16 日上午，国新办举行了第三次全国经济普查新闻发布会。国务院第三次全国经济普查领导小组副组长、国家统计局局长马建堂介绍了第三次全国经济普查的主要目的。

马建堂指出，第三次全国经济普查的主要目的是全面调查了解我国第二产业和第三产业的发展规模及布局，进一步查实服务业、战略性新兴产业和小微企业的发展状况，摸清我国各类单位的基本情况，全面更新覆盖国民经济各行业的基本单位名录库、基础信息数据库和统计电子地理信息系统，为加强和改善宏观调控、加快经济结构战略性调整、科学制定中长期发展规划，提供科学、准确的统计信息支持。

二、调查对象与调查单位

调查对象即统计总体，是根据调查目的所确定的研究事物的总体或调查范围。统计总体这一概念在统计调查阶段称为调查对象。调查对象是根据调查目的确定的调查研究的总体。在确定调查对象时，必须确定调查单位和报告单位。调查单位就是总体单位，也是调查对象的组成要素，即调查对象所包含的具体单位，是搜集数据资料的基本单位。

小贴士

确定调查对象和调查单位时应注意的问题

第一，调查单位与填报单位是有区别的，调查单位是调查项目的承担者，而填报单位是调查中填报调查资料的单位。例如，对某地区物流运输企业设备进行普查，调查单位为该地区物流运输企业的每台设备，而填报单位是该地区每家物流运输企业。但在有些情况下两者又是一致的。例如，在进行职工基本情况调查时，调查单位和填报单位都是每一个职工。在调查方案设计中，当两者不一致时，应当明确从何处取得资料并防止调查单位重复和遗漏。

第二，不同的调查方式会产生不同的调查单位。普查时，调查总体内所包括的全部单位都是调查单位；重点调查时，只有选定的少数重点单位是调查单位；典型调查时，只有选出的有代表性的单位是调查单位；抽样调查时，用各种抽样方法抽出的样本单位是调查单位。

三、调查项目

调查项目是调查的具体内容，它可以是调查单位的数量特征。例如，一个人的年龄、收入，一个企业的职工人数、产值；也可以是调查单位的某种属性特征，如一个人的性别、职业，一个企业的经济类型等。

确定调查项目时，首先，应注意所选择的项目能够取得确切资料；其次，注意所选择的项目应有确切的含义和统一解释；再次，还要注意各项目之间的联系和衔接，便于核对和分析。

四、调查表

调查表是用于登记调查数据的一种表格，其目的是保证统计资料的规范化和标准化。调查表有单一表和一览表两种形式，单一表是一个调查单位填写一份表格，可以容纳较多的项目。一览表是许多调查单位共同填写一份表格，在调查项目不多时较为简便，且便于合计和核对差错。为了正确填写调查表，须附有填表说明和项目解释。

调查表一般由表头、表体和表外附加三部分组成，如表 2-1 所示。

表 2-1　调查表组成部分

表头	某企业基本情况报表　　制表机关：××××			
表体	调查项目	项目代码	计量单位	实际数值
	甲	1	人	
	乙	2	台	
	丙	3	人	
	…	…	…	…
表外附加	说明：……			

五、调查时间和调查时限

为确保调查资料的准确性，统计调查必须规定调查时间和调查时限。

调查时间指调查资料所属时间。如果调查的是时期现象，调查时间是资料所反映的起讫时间；如果调查的是时点现象，调查时间是统一规定的标准时点。

调查时限是进行调查工作的期限，包括搜集资料和报送资料的整个工作所需要的时间。如某上级部门要求所属企业在 2017 年 1 月 10 日上报 2016 年产成品库存资料，则调查时间是标准时间 2016 年 12 月 31 日，调查时限是 10 天。再如，第四次人口普查规定的标准时间是“2000 年 11 月 1 日 0 时”，并要求在“2000 年 11 月 1 日至 10 日完成普查的登记工作”。2000 年 11 月 1 日 0 时就是调查时间，2000 年 11 月 1 日至 10 日就是调查时限。

六、调查的组织工作

调查组织实施是指从组织上保证调查工作顺利进行的措施。其主要内容包括：调查工作的组织机构；调查人员的配备；调查方式的选择；调查工作地点的确定；调查文件的准备；经费来源和主要开支计划，等等。对于规模大又缺乏经验的统计调查，事先需要进行试点调查。

案例学习

大学校园旧货交易市场调查方案

每当毕业季临近，大学生在大学校园里处理自己不需要的东西时，往往都采用当作废品变卖、送给自己朋友等低价值转换方法。据观察，现在大学校园除了毕业生临时交易市场和平时采用张贴小广告等松散形式进行交易以外，别无其他处理旧货的市场，而这种临时性交易往往具有时间地点不固定、缺乏信用、破坏校园环境等缺点。

因此，建立一个大学生旧货交易平台便具有可行性。为此我们以华北水利水电大学及其周边几所大学为重点，就建立大学生旧货交易平台所面对的市场前景进行调查。

一、调查目的

1. 通过这次调查，为以后在大学校园建立旧货交易市场提供第一手资料。

2. 宣传旧货交易相关知识。

3. 与具有合作意愿的同学建立起初步联系。

二、调查对象

本次调查对象为本校及周边学校的在校大学生和大学周边旧货市场店主。

三、调查内容

1. 了解大学生对旧货交易的态度。

2. 了解大学生处理旧货的具体方式。

3. 了解大学周边旧货市场运作形式及规模。

4. 了解大学校园旧货市场状况。

5. 了解大学生对旧货交易形式的期望。

6. 受访者的背景资料。

四、调查方式

采用随机抽样调查方式。

五、调查方法

1. 文献调查，利用互联网、图书馆进行相关资料的收集、整理与分析。

2. 问卷调查，利用设计的问卷对大学校园相关人员进行调查。

3. 询问调查和观察调查。

六、调查时间安排

计划利用学期开学初的一个半月时间完成此项方案，具体安排如下：

文献查阅——2 天；

问卷设计——3 天；

问卷、询问和观察——30 天；

数据资料处理——3 天；

报撰写统计分析报告——7 天。

七、调查经费预算

预算费用（问卷印刷费、调查实施费、报告制作费等）共计约 1000 元。

八、调查工作的组织实施

此项调查由华北水利水电大学团委指导并负责调查经费的筹集，由校学生会具体负责组织实施，勤工俭学协会协助，参与调查人员由各个院系推荐。

第三节　统计调查方法

统计调查方法指的是搜集调查对象原始资料的方法，也就是调查者向被调查者搜集答案的方法。我国统计调查工作要求建立以必要的周期性普查为基础、经常性的抽样调查为主体，同时辅之以重点调查、科学推算和少量的全面报表综合运用的统计调查方法体系。统计调查方法按组织方式可以分为统计报表和专门调查，专门调查包括普查、抽样调查、重点调查、典型调查等。

一、统计报表

（一）统计报表的概念

统计报表是企业、事业、行政单位、各级政府机构和各业务主管部门，依照《中华人民共和国统计法》的有关规定，按照统一的表式、报送程序、报送时间等统一要求，自上而下布置，自下而上逐级上报提供统计资料的一种统计调查方式。我国大多数统计报表要求调查

对象全部单位填报，属于全面调查范畴。

统计报表制度是一国掌握社会经济发展动态的基本情况、稳定获取统计资料的重要途径。国家统计部门利用统计报表为政府制定、编制和检查计划提供了大量可靠的依据，也可以指导有关部门、企业的生产、经营、管理和决策。



小贴士

统 计 报 表

统计报表是按国家规定的内容、表式和时间程序，自上而下统一布置、自下而上逐级提供基本统计资料的一种调查方式，分定期报表和不定期报表两种形式，二者都是我国统计调查的基本组织形式。

在我国，统计报表中绝大部分是定期统计报表，统计报表所包含的范围比较全面，项目比较系统，分组也比较齐全，指标的内容和调查周期相对稳定。

（二）统计报表的特点

1. 统计报表是国家和地方政府部门收集统计数据的主要方式

统计报表在我国几十年的政府统计工作中已形成了一套比较系统、完备的统计报表制度，是国家和地方政府部门收集统计数据的主要方式。

2. 统计报表的资料来源于基层单位的原始记录

从原始记录到统计报表，中间会经过统计台账和企业内部报表。

3. 统计报表只是一种补充性的统计调查

统计报表较笨重，缺乏灵活性，而且财力、人力投入较大；调查效益差，基层负担重，中间环节多，容易受到行政干扰，数据信息质量很难得到保证。面对日益发展的多种经济成分、多种分配方式等复杂多变的调查对象，它只能作为一种补充性的统计调查方法。

（三）统计报表的种类

1. 按调查对象范围的不同，可以分为全面统计报表和非全面统计报表

全面统计报表要求调查对象的每个单位都填报，非全面统计报表只要求调查对象的一部分单位填报。我国大多数统计报表要求调查对象全部单位填报，属于全面调查范畴，所以又称全面统计报表。

2. 按报送周期长短的不同，可以分为日报、旬报、月报、季报、半年报、年报

日报、旬报可以及时反映经济发展的状况，只限于少量最主要的指标。月报、季报主要用于检查月、季计划执行情况，半年报、年报是具有总结性的报表，指标项目最多，内容全面完整，对摸清国民经济运行的详细情况有重要作用。

3. 按报送方式的不同，可以分为邮寄报表和电讯报表

电讯报表包括电报、电话、传真和网上报表等方式。

4. 按填报单位的不同，可以分为基层报表和综合报表

基层报表主要由基层企事业单位填报，它所提供的原始资料是统计的基础资料。综合报表是由主管部门根据基层报表逐级汇总填报的统计报表。

5. 按内容和实施范围的不同，可以分为国家报表、业务部门报表和地方报表

目前我国统计报表是由国家统计报表、业务部门统计报表和地方统计报表组成的，其中国家统计报表是统计报表体系的基本部分。

国家统计报表是在全国范围内实施的，用以搜集有关社会经济发展全貌的基本统计资料。部门统计报表是在各个部门系统内实施的，用以搜集各个部门自身发展情况的专业统计资料。地方报表是在各地所辖范围内实施的，用以搜集满足地方所需要的区域统计资料。业务部门报表和地方报表都是国家统计报表的有效补充。

（四）统计报表的基本内容

1. 报表目录

报表目录是为了让填报单位了解在什么时间、用什么方式、向什么单位报送什么报表，具体包括统计报表名称、填报单位、调查对象的统计范围、报送时间、报送程序。

2. 报表表式

它是统计报表制度的主体，是指统计报表的具体格式，包括表名、表号、报表期限、填报单位、报出时间、报送方式、主栏项目、宾栏指标、表下补充资料、填报单位负责人和填表人签章等。

3. 填表说明

它是对统计报表的统计范围、指标等做出的规定。

（1）填报范围

即统计报表的范围，规定每种统计报表的填报单位、各级统计部门与主管的范围等。

（2）指标解释

对列入表统计指标的口径、计算方法及其他问题的具体说明。

（3）分类目标

统计报表主栏中应该进行填报的有关项目的分类。

（4）其他有关事项的规定

除了以上各项规定以外的一些注意事项，如报送日期、报送方式、报送份数等。

（五）统计报表的资料来源

统计报表的资料来源主要是基层的原始记录、统计台账及基层企业内部报表。

原始记录是基层单位通过一定的表格形式，对其生产、经营、管理活动的具体内容和状

况进行的最初的数字和文字记载。例如企业产品产量记录、职工出勤记录、产品入库单等。原始记录具有广泛性、经常性、群众性和真实性的特点。

统计台账是基层单位根据统计报表要求和基层经营管理需要而按时间顺序设置的一种系统积累统计资料的表册。统计台账有单指标和多指标的统计台账，可以作为填报统计报表的依据，又可以积累统计资料，是积累历史资料的手段和工具。

基层内部报表只适合本企业，它是对本企业内部的生产、经营活动情况的填报。

二、普查

（一）普查的概念

普查是为某一特定的目的而专门组织的一次性全面调查。普查主要是调查一定时点的社会经济现象的总量，搜集那些不能够或者不适宜用定期全面报表搜集的统计资料。普查一般用来调查社会经济现象在一定时点上的具体数量情况、基本面貌，如人口普查、经济普查等。普查范围广，搜集的资料全面，能提供较多、较详细的原始资料，是摸清国情国力、为国家制定有关政策或措施提供依据的重要手段。

普查和全面统计报表都属于全面调查，但是二者并不能互相代替。普查属于不连续调查，调查内容主要是反映国情国力方面的基本统计资料；而全面统计报表属于连续调查，调查内容主要是需要经常掌握的各种统计资料。全面统计报表要经常填报，因此报表内容固定，调查项目较少；而普查是专门组织的一次性调查，在调查时可以包括更多的单位，分组更细、项目更多。

因此，有些社会经济现象不可能也不需要进行经常性调查，但又需要掌握比较全面、详细的资料时，就可通过普查来解决。普查花费的人力、物力和时间较多，不宜经常组织，取得经常性的统计资料还要靠全面统计报表。

（二）普查的特点

1. 普查具有周期性

由于人力、财力和时间所限，普查不可能年年搞，更不能事事搞，而只能对主要的调查项目每隔若干年进行一次。通常每隔 5 年或 10 年进行一次，即周期性地进行普查。目前，政府部门周期性的普查项目有人口普查、农业普查和经济普查等。

2. 普查需要规定标准调查时间

普查一般需要规定统一的标准调查时间，以避免调查数据的重复和遗漏，保证普查结果的准确性。例如，我国前四次人口普查的标准时点定为普查年份的 7 月 1 日零时；第五次全国人口普查的标准时点定为 2000 年 11 月 1 日零时；2008 年经济普查所定的标准时点为 2008 年 12 月 31 日。标准时间一般定在调查对象比较集中、变动相对较小的时段。

3. 普查需要规定统一的普查项目和指标

普查必须按照统一规定的项目和指标进行登记，不得任意改变或增减，以便综合汇总。对同一现象的普查，每次普查的项目和指标力求一致，便于与历次普查资料进行对比分析。

4. 普查由专门机构来组织

普查的组织方式基本上有两种：一是组织专门的普查机构，配备一定数量的普查人员，对被调查者直接进行登记；二是利用被调查者的原始记录和核算资料，发放一定的调查表格由被调查者填报。对于后一种方式，仍需组织一定的普查机构、配备一定的专门人员，以对整个普查工作进行组织领导。

（三）普查的种类

1. 按组织形式分类

普查的组织方式有两种：一是通过专门组织的普查机构，配备一定数量的人员，对调查单位直接进行登记，如人口普查；二是利用企事业单位日常核算资料和报表资料，发放一定的调查表格，由调查单位填报，如物资库存普查、快速普查等。

2. 按汇总资料特点分类

按汇总资料的特点不同，可以分为一般普查和快速普查。一般普查中逐级上报资料，快速普查中越过中间环节，由基层单位将资料直接报送给最高领导机关。



小贴士

几种重要普查的周期

根据《中华人民共和国统计法》的规定和国务院关于建立国家普查制度、改革统计调查体系的要求，国家统计局于 1994 年正式建立了周期性的普查制度。普查项目包括人口、农业、工业、第三产业和基本单位。人口普查、第三产业普查、工业普查、农业普查每 10 年进行一次，分别在逢 0、3、5、7 的年份实施；基本单位普查每 5 年进行一次，在逢 1、6 的年份实施。

2003 年国家对普查项目和周期安排重新做了调整，我国在 10 年中，也就是 2004 年到 2014 年，开展 3 项 4 次普查，即：逢尾数为 0 的年份做人口普查，10 年中进行一次；逢尾数为 6 的年份做农业普查，10 年中进行一次；逢尾数为 3 和 8 的年份做经济普查，10 年中进行两次。

经济普查实际上是将过去的工业普查、第三产业普查和基本单位普查合并，再加上过去没有纳入普查范围的建筑业合起来一起去做，叫经济普查。现在我国重要的国情国力普查已形成制度。

三、抽样调查

（一）抽样调查的概念

抽样调查是按随机原则从总体中选取一部分单位进行观察，用以推算总体数量的一种非全面调查。它是根据概率理论从总体中随机抽取一部分单位进行研究，其目的是推断总体的特征，它属于非全面调查。

社会科学研究经常使用抽样调查。例如，政府机构想了解全市下岗职工的年龄构成及再就业状况。此时的总体应是全市所有下岗人员，要想得知下岗人员的年龄构成及再就业状况，就应做一次全面调查。然而总体如此庞大，鉴于人力、物力、财力与时间等方面的限制，全面调查显然是不现实的，而科学的做法是运用抽样技术选择一个好样本，用样本信息对整个总体做某些结论。



小贴士

抽 样 调 查

抽样调查是非全面调查中最具计量科学依据的方式。其科学价值在于既能省时、省力，又能较准确地达到认识总体数量特征的目的，并且抽样误差可以计算和控制。因此，我国统计调查方法改革的目标模式是：建立一个以必要的周期性普查为基础、以经常性的抽样调查为主体，同时辅之以重点调查、科学推算和全面报表综合运用的统计方法体系。

（二）抽样调查的特点

1. 经济性好

抽样调查虽然是非全面调查，但它能达到对总体数量特征的认识，它只调查总体中的一部分单位，工作量小，投入少，因而可以节省大量的人力、物力、财力和时间，能够以较少的投入取得必要的统计数据。抽样调查是非全面调查中最完善、最有科学根据的方式方法。

2. 时效性强

由于抽样调查的工作量比普查工作量小，调查时间、数据处理的时间较短，从而提高了数据的时效性。

3. 准确性高

采用抽样调查，可以避免或减少层层汇总上报中的行政干预行为，取得的统计数据比较准确；另外，抽样调查覆盖面小，工作量相对较小，可使各环节的工作做得更细致，误差较小。当然，用样本数据去推断总体，不可避免地会产生抽样误差，但这种误差的大小是可以计算并加以控制的。

4. 适应性广

抽样调查不仅比普查划算，而且应用范围很广泛。例如，要测试灯泡的使用寿命、轮胎的里程、新研制炮弹的射程等，必须采用抽样调查；又如，居民家庭收支状况、人口变动情况、电视收视率等也适合用抽样调查。抽样调查已成为世界各国普遍采用的一种调查类型。

（三）抽样调查的作用

①能够解决全面调查无法或难以解决的问题。
②可以补充和订正全面调查的结果。
③可用于生产过程中产品质量的检查和控制。
④可用于对总体的某种假设进行检验。

（四）抽样调查的基本形式

抽样调查的基本形式有简单随机抽样、类型随机抽样、等距抽样、整群抽样和多阶段抽样。

1. 简单随机抽样

按随机原则直接从总体中抽选样本单位进行调查。简单随机抽样通常用抽签的方式抽取所要调查的单位。

2. 类型随机抽样

先对总体各单位按主要的标志加以分类，然后按随机原则从各类中抽取一定单位数进行调查。

3. 等距抽样

将总体各单位按某一标志大小顺序排列，然后依一定间隔抽取样本单位进行调查。

4. 整群抽样

整群抽样中，先将总体单位划分为许多群，然后以群为单位，从其中随机抽取部分群，再对中选群的所有单位进行全面调查。

5. 多阶段抽样

多阶段抽样，就是把抽取样本单位的过程分为两个以上的阶段进行。先从总体中抽选若干大的样本单位，这是第一阶段抽样。然后从被抽中若干大的单位中抽选较小的样本单位，这是第二阶段抽样。以此类推，再在被抽中的单位中抽选较小的样本单位，还可以有第三阶段、第四阶段等，最后推出所需的样本来。

例如，某地区教委要了解小学生的健康状况，直接抽查有一定的困难，因此可以采取多阶段抽样。首先按照随机原则抽选一些学校，然后在被抽中的学校中随机抽选一些班级，再在被抽到的班级中抽选一些学生进行调查。

（五）抽样调查适用的场合

①有些现象不可能进行全面调查，如测试灯泡的使用寿命、轮胎的里程、新研制炮弹的射程等必须采用抽样调查。

②有些现象不必进行全面调查。

③有些现象来不及进行全面调查。

④全面调查以后，采取抽样调查可以对全面调查资料进行修正补充。

（六）抽样调查与普查的关系

普查和抽样调查都很常用。普查具有周期性，一般 5 年或 10 年进行一次；抽样调查主要用于常规性调查，包括年度、季度和月度统计调查。我国的统计法明确规定，统计调查以周期性普查为基础，以经常性抽样调查为主体。由此不难看出，普查和抽样调查在统计调查中都具有十分重要的地位，是统计调查的两大支柱。

1. 抽样调查与普查的比较

抽样调查和普查同时存在并得以广泛应用的现实足以说明二者各有所长、各有不足，彼此不可替代，如表 2-2 所示。

表 2-2　抽样调查与普查的比较

比较的内容	抽样调查	普查
1.省费用	优	
2.及时性	优	
3.灵活性	优	
4.详细分组资料和满足分级管理		优
5.准确性或可靠性	不确定	
（1）抽样误差	存在	不存在
（2）非抽样误差	存在	存在

表 2-2 从五个方面对抽样调查和普查的优劣进行比较。从中可以看到，在费用、时间和灵活性三个方面，抽样调查具有优势，因为抽样调查只抽取部分样本进行调查，工作量远远小于普查。

2. 抽样调查与普查互为补充

抽样调查和普查并存，相互补充，构成了一个完整的统计调查体系。

（1）普查可以为抽样调查提供比较全面、详细的抽样框

抽样离不开抽样框，而普查的详细资料本身就是很好的抽样框。从这个意义上讲，普查是抽样调查的基础。

（2）抽样调查是对普查数据进行质量检查的有效工具

利用抽样调查对普查质量进行检验，目前已广泛地应用于各国的普查过程。

（3）两种方法可以结合使用

对于一些重大的国情国力问题，如人口、农业等，有必要动用大量的财力和物力进行普查，以便得到非常详细的资料。然而，普查耗资大、耗时长，不可能也没有必要经常进行。在普查的间隔年度及季度和月度等进行的经常性调查，应当通过抽样调查来完成。

在普查中可以对一些规模很小的单位不进行逐个调查，而是通过抽样调查获取资料。例如，我国 2004 年进行的经济普查对个体户采用的就是抽样调查。

四、重点调查

（一）重点调查的概念

重点调查是专门组织的一种非全面调查，它是对所要调查的全部单位选择一部分重点单位进行调查。重点调查的关键是选择好重点单位。所谓重点单位，是从标志量的方面而言的，尽管这些单位在全部单位中只是一部分，但这些单位的某一主要标志量占总体单位标志总量的绝大比重。对这些单位进行调查，就可以了解调查对象的基本情况。

例如，要了解全国钢铁企业的生产状况，只要选择鞍钢、宝钢、首钢、太钢、包钢、武钢等大型钢铁企业作为重点单位进行调查就可以了。因为这些大型钢铁企业的产量占全国钢铁产量的绝大比重。对这部分重点单位进行调查所取得的数据一般能够反映总体的基本情况。

当然，重点单位除了可能是企业外，也可能是一些地区、城市。例如，为了了解黄金周旅游市场的基本情况，对全国几个主要旅游景区进行调查就可以了。重点调查选择的调查单位少，因此具有省时省力的优点，但只能取得反映总体基本情况的资料。由于重点单位与一般单位差别很大，因此，重点调查不可推断总体数量特征。

（二）重点调查的特点

1. 重点调查既经济又快速

重点调查是范围比较小的非全面调查，所投入的人力、物力少，又可以较快地收集到统计信息资料。

2. 重点调查中重点单位的选择具有客观性

重点调查中重点单位的选择着眼于标志量的比重，因而重点单位的选择具有客观性。重点调查可以定期进行，也可以不定期进行，重点调查实际上是范围比较小的全面调查，它的目的是反映现象总体的基本情况。

3. 重点单位对于总体来说最具代表性，但不能推断总体

当调查目的是掌握现象的基本情况，而部分单位又能比较集中地反映所研究的项目和指标时，可使用重点调查法。另外，重点单位会随时间、地点、研究内容的不同而发生变化。

（三）重点调查与抽样调查的区别和联系

重点调查与抽样调查的联系：都是专门组织的非全面调查，具有调查单位少、省时省力

的特点，在选取调查单位时不受主观因素的影响。

但二者之间有明显的区别，具体如下。

首先，调查单位的意义和取得方式不同。重点调查是选择为数不多但标志量占总体标志总量绝大比重的单位进行调查；抽样调查中的样本单位是按照随机原则从研究总体中抽取的，具有较高的代表性。

其次，二者的研究目的不同。重点调查是为了解现象总体的基本情况，但不能推断总体总量；抽样调查的目的在于以样本量来推断总体总量。

最后，适用场合不同。重点调查适用于部分单位能比较集中地反映所研究的项目或指标的场合；抽样调查最适合于不能或很难进行全面调查而又需要全面数值的场合，在能进行全面调查的场合也有独到的作用。

（四）重点单位的选择

根据调查的目的和任务，选择管理水平较高、核算制度比较健全、统计力量比较充实、统计基础比较巩固、资料容易取得且质量较高的单位。

五、典型调查

（一）典型调查的概念

典型调查是根据调查的任务和目的，在对所研究的现象总体进行初步分析的基础上，有意识地选择若干具有代表性的单位进行调查，借以认识事物发展变化的规律。典型调查在社会学研究中经常使用，也是一种非全面调查。

（二）典型调查的特点

1. 典型调查具有深入细致性

由于调查单位少，而且调查单位经过分析选择，具有典型性，因此便于进行深入、具体、周密的调查，能够发现丰富生动的信息，适用于对新情况、新问题的调研。典型调查具有深入、细致性的优点，但也有不能推断总体的缺点。它既可以搜集数字资料，又可以搜集不能用数字反映的实际情况。

2. 取决于调查者的主观判断和决策

典型调查，调查单位更多地取决于调查者的主观判断和决策。典型调查是对所研究的现象总体进行初步分析的基础上，有意识地选择若干具有代表性的单位进行调查，因此它更多地取决于调查者的主观判断和决策。

3. 投入少、效率高

节约人力、物力和时间，具有机动灵活性，可以提高调查资料的效率和时效性。

（三）典型调查的组织方式

组织典型调查的关键在于如何选择典型单位，应根据调查目的和研究主题确定，通过比较筛选实施，通常有以下两种做法。

1. “解剖麻雀”式

“解剖麻雀”式是对个别典型单位进行的调查研究，通过“解剖麻雀”，掌握总体内部结构和发展变化的规律，达到对现象总体基本情况的认识。如果调查总体各单位的差异较小，解剖一个或几个“麻雀”就行了。如果典型调查的目的是了解总体发展的一般情况，则可以选择总体中的中等水平单位作为典型单位。

2. 划类选典式

划类选典式是先对现象总体按某主要标志划类，再在各类中选择典型单位予以调查。

如果典型调查的目的是了解总体发展的基本情况，且调查总体各单位的差异较大，则需要把调查总体按照一定的标准划分成几种类型，然后在各类型中选取典型单位进行调查，以认识各单位的一般水平和估计总体大致水平。

（四）典型调查与重点调查的区别

典型调查单位的选择取决于调查者的主观判断，重点调查单位的选择具有客观性。典型调查在一定条件下可以用典型单位的量推断总体总量，但无法估计误差，推断结果只是一个近似值；重点调查不具备用重点单位的量推断总体总量的条件。

抽样调查和重点调查、典型调查的根本区别在于选取调查单位的方法不同。

六、各种调查方式的结合运用

不同的统计调查方法各有其特点和作用。在实际工作中，并非单用一种方式方法，而是多种方式方法的结合运用。这是因为：国民经济和社会发展情况复杂，国民经济门类众多，必须应用多种多样的统计调查方法，才能搜集到丰富的统计资料；任何一种统计调查方法都有它的优越性与局限性，各有不同的实施条件，只用一种统计调查方法，不能满足多种需要。

第四节　统计调查报告

一、统计调查报告概述

（一）统计调查报告的概念

统计调查报告是在统计调查的基础上，对所研究的社会经济现象进行大量观察之后，

选择有代表性的单位或地区进行深入调查研究以反映是非曲直及其本质和规律的一种专业应用文。

（二）统计调查报告的特点

1. 取材直接

统计分析报告主要使用反映总体情况的统计资料，辅以实际调查而获取的具体材料。而统计调查报告则主要使用直接调查的具体材料。这些材料主要是从抽样调查、重点调查和典型调查等非全面的专门调查中得来的。

要求调查人员必须深入实际，亲自调查采集第一手资料，掌握实际情况。如果调查人员没有直接参加实际调查，而主要依靠间接材料，就很难写出报告，即使勉强写出来了也不符合要求。

2. 内容深入

统计调查报告通过非全面的专门调查来反映部分单位的社会经济情况，并不要求直接反映总体情况。由于调查单位较少，调查范围较窄，因此可以深入地进行调查研究，细致地解剖社会经济现象，尽力挖掘其客观意义和规律特征，从而使统计调查报告比其他统计应用文内容更深入、材料更具体、分析更透彻，更能反映客观事物发展变化全过程。

3. 事例典型

统计调查报告要让事实说话，用事实来说明观点。这不仅要用较多的统计数字和综合材料，而且要用大量的典型事例，有的甚至是一个完整的典型事件贯穿全文。

由于典型事实具有一定的代表性，因此通过对它们进行深入的研究，在一定条件下可以认识同类事物的共同本质和一般规律。

4. 叙议结合

统计调查报告的陈述性和说明性较强，因此主要表达方式是叙述和议论，又以叙述为主。统计调查报告的陈述既要简洁而概括，又要具体而生动。其他统计应用文也有对事实的叙述，但只取事实的片断，不求系统的完整，而统计调查报告，尤其是反映典型事件和新事物的调查报告，要求展示事物发生、发展的过程，因而具有一定的完整性。

统计调查报告不仅要反映事实的经过，还要对事实进行分析和议论，找出规律性的东西。但是必须以事实为基础和依据，绝不能脱离事实去作空洞的议论。

（三）统计调查报告的种类

1. 按内容性质分类

（1）反映一般情况的调查报告

这种调查报告研究某一社会经济现象的基本情况，为领导掌握情况、解决问题、制定政策提供参考和依据。这种调查报告既可以反映成绩，也可以反映问题，还可以预测未来。

（2）研究政策的调查报告

这种调查报告对党和政府的方针、政策在贯彻执行中的各种情况进行反馈。通过实践的

检验，对其不完善的地方提出改进的意见，以便修正和调整，使之更加符合客观实际。这种调查报告中分析和议论的成分较重。

（3）总结经验的调查报告

这种调查报告主要是推广在实践工作中比较成熟而具有指导意义的经验，以先进典型带动一般。写好这种调查报告要善于发现、挖掘和总结典型事物，介绍的做法和经验要明确具体、具有可操作性。

（4）揭露问题的调查报告

这种调查报告针对阻碍社会经济发展的带有倾向性的问题或事件，用确凿的事实揭露矛盾，反映问题的严重程度和危害性，分析问题产生的原因，找出解决问题的办法，从而达到澄清事实、明辨是非、吸取教训的目的。

（5）反映新事物的调查报告

这种调查报告比较完整、全面地反映某一新事物产生的背景、发展的过程及存在的问题，研究它的性质、意义和规律，以便其正常成长和及时推广。

2. 按内容范围分类

（1）专题调查报告

专题调查报告是对某一个或某一方面的问题进行专题调查而写成的报告。它反映的对象可以是个体，即针对一个专题对一个单位进行调查；也可以是群体，即针对一个专题对部分单位进行调查。专题调查报告选题灵活、内容集中、针对性强、易于较快拿出成果。专题调查报告具有一定的深度，能为领导解决某方面的问题、制定某方面政策提供依据。

（2）综合调查报告

综合调查报告是对一个单位、一个地区多方面情况进行全面调查而写成的报告。要求把调查对象各方面的情况联系起来进行综合研究，进行全面评价。综合调查报告主要用于为领导从全局上指导工作提供决策咨询。

3. 按功能目的分类

（1）描述性调查报告

这是以了解某一事件或某一问题的全貌或发展过程为目的的调查报告。它要解决“是什么”的问题，如民意调查、国情国力调查、基本情况调查等。这种调查报告以反映情况为主，描述事实，摸清问题，不必进行过多的议论。

（2）解释性调查报告

这是以探讨社会经济现象之间的因果关系为目的而写的调查报告，它要解决“为什么”的问题，要求对调查对象进行深入研究，探讨原因，找出规律。因此，分析、议论较多。

（3）预测性调查报告

它是在调查事物的历史和现状的规律，对未来的发展方向进行判断和描述，为领导制定规划和决策提供方向性的参考意见。预测性调查报告要解决的是“将会怎么样”的问题。

（四）统计调查报告的作用

统计调查报告可以为国家制定方针、政策及各部门、企事业单位的领导者做出正确决策提供依据；调查报告可以通过典型调查、宣传介绍先进经验和先进人物事迹，借以指导全面工作；调查报告可以通过典型调查揭露社会问题，鞭挞不良倾向，改正工作中的失误；调查报告可以通过调查揭露事实真相，说明和回答社会问题。

二、统计调查报告的撰写

（一）撰写调查报告的重要性

调查报告是调查人员对某种事物或某个问题进行深入细致的调查后，经过认真分析研究而形成的一种报告形式。书面调查报告的功能体现在以下三点。

1. 调查报告是调查工作的最终成果

我们知道，调查活动是一个有始有终的活动，从制订调查方案、搜集资料、加工整理和分析研究到撰写并提交调查报告，是一个完整的工作程序，缺一不可。所以，调查报告是调查成果的集中体现。

2. 调查报告是从感性认识到理性认识飞跃过程的反映

调查报告比起调查资料来，更便于阅读和理解，它能把死数字变成活情况，起到透过现象看本质的作用，使感性认识上升为理性认识，便于更好地指导实践活动。

3. 调查报告是为各部门管理者、为社会、为企业服务的重要形式

一份好的调查报告能对各项活动提供有效的导向作用，同时对各部门管理者在了解情况、分析问题、制定决策和编制计划及控制、协调、监督等各方面起到积极的作用。

要撰写好一份出色的调查报告，必须了解调查报告的特点，掌握整个调查报告的撰写步骤、报告的结构和撰写方法，使调查报告在实际工作或理论研究中发挥应有的作用。

（二）撰写统计调查报告的基本要求

1. 实事求是

调查报告作为调查研究的成果，最基本的特点就是尊重客观实际，用事实说话。

2. 符合经济规律及有关政策的规定

经济活动有其特有的规律性。因而，要摸清市场规律，及时掌握市场规律的变化，研究其变化的原因，加深对市场规律的认识，同时密切注意各个时期党和国家有关方针政策的变化，可以使问题剖析得更加透彻，使调查报告更加真实、准确。

3. 观点与数据要结合运用

调查报告的独特风格就是以调查资料为依据，而资料中的数据资料显得尤为重要，数据

资料具有很强的概括力和表现力。用数据证明事实的真相往往比长篇大论更能使人信服。在调查中，常常会碰到某些问题、观点用很多叙述都难以表达清楚，而用一个数字就可以使事物的全貌一目了然。

此外，调查报告还必须做到主题突出、结构严谨、条理清楚、文字简洁，同时，尽量用表格、图表、照片或其他可视物品来补充正文中的关键信息。直观可视的图表等对帮助报告撰写人和读者之间进行交流很有好处，也可以增强报告的明了程度和效果。

（三）书面调查报告的格式

调查报告一般是由标题、导语、正文（主体）、结尾等几部分组成的。

1. 标题

标题应是画龙点睛之笔，它必须准确揭示调查报告的主题思想，做到题文相符。标题要简单明了、高度概括、具有较强的吸引力。

标题的写法灵活多样，一般有两种：单标题和双标题。

（1）单标题

①一般文章的标题形式。

有的点明基本观点，如《应重视农村水利设施建设》；有的反映内容范围，如《我市青年结婚费用支出情况》；有的提出问题，如《小氮肥厂的出路何在》等。

②公文的标题形式。

有的像公文标题一样，由内容范围和文体种类两部分组成，常用“关于……的调查”或“关于……的调查报告”来表示，如《关于当前物价上涨情况的调查报告》；有的还可以加上编写调查报告的单位，如《中国人民银行××县支行关于货币流通量的调查报告》。

（2）双标题

双标题分为正标题和副标题。正标题往往是“虚题”，揭示调查报告的基本观点。副标题往往是“实题”，补充说明调查的对象、内容、范围等情况。例如，《提高农民工的科学文化水平是致富的重要条件——60 个不同文化水平农民工工作生活差距的调查》。

2. 导语

调查报告的导语又叫开头或前言。通过调查报告的引言可以使阅读者对报告有所了解，调查报告的引言通常包括调查背景和调查目的两个部分。在调查背景部分的报告内容中应对调查的由来或受委托进行该项调查的原因做出说明，对调查项目的必要性进行简要的解释。常见的导语有叙述式、论述式和提问式三种类型。

3. 主体

主体是统计调查报告的主要部分。主体的任务是用事实反映调查研究的结果，用材料阐明调查报告的观点，使观点与材料相统一。主体的写作必须做到以下几点：

①观点要明确；

②材料要充分；

③分析要具体；

④结构要合理。

4. 结尾

统计调查报告的结尾是总结全文、得出结论、深化主题、解决问题的部分。结尾的写法没有固定不变的形式，要根据写作目的和主体部分的内容灵活选择。常用的结尾方式有指明方向、提出建议、总结全文等几种。

（四）撰写调查报告的技巧

撰写调查报告的技巧主要包括表达的技巧、图表运用的技巧等。其中表达的技巧又包括叙述的技巧、说明的技巧、议论的技巧、语言的技巧等。

1. 叙述的技巧

叙述主要用于市场调查报告的开头部分，叙述事情的来龙去脉，表明调查的目的和根据、调查的过程和结果。调查报告常用的叙述技巧有：概括叙述、按时间顺序叙述。

①概括叙述：将调查过程和情况概略地陈述，不需要对事件的细节加以说明，这是一种“浓缩型”的快节奏叙述，文字简约、一带而过，给人整体、全面的认识，以适合市场调查报告快速及时反映市场变化的需要。

②按时间顺序叙述：在叙述调查的目的、对象、经过时，往往采用按时间顺序叙述的方法，次序井然，前后连贯。

2. 说明的技巧

市场调查报告常用的说明技巧有：数字说明、分类说明、对比说明和举例说明。

（1）数字说明

反映市场发展变化情况的市场调查报告需要运用大量的数据，以增强调查报告的精确性和可信度。

（2）分类说明

市场调查中所获资料杂乱无章，根据主要表达的需要，可将调查数据资料按一定标准分为几类，分别说明。

（3）对比说明

市场调查报告中有关情况、数字说明，往往采用对比形式，以便全面深入地反映市场变化的情况。对比要清楚事务的可比性，在同标准的前提下，做切合实际的比较。

（4）举例说明

为说明市场发展变化情况，举出具体典型实例，这也是常用的方法。市场调查中，会遇到大量实例，应从中选取有代表性的例子。

3. 议论的技巧

市场调查报告常用的议论技巧有：归纳论证和局部论证。

①归纳论证：对大量材料进行归纳分析研究，得出共性结论，从而形成论证。这一过程主要运用议论方式，所得结论是从具体事实中归纳出来的。

②局部论证：市场调查报告不同于议论文，不可能全篇论证，只是在情况分析和对未来预测中进行局部论证。

4. 语言运用的技巧

语言运用的技巧主要包括用词方面的技巧。在用词方面，市场调查报告中数词用得较多，因为市场调查离不开数字，很多问题要用数字说明。可以说，数词在市场调查报告中以其特有的优势越来越显示出重要作用。此外，还经常使用专业词汇，以反映市场的发展变化，如“商品流通”、“市场竞争”等词。为使语言表达准确，调查报告撰写者还需要熟悉有关专业术语。

5. 图表运用的技巧

在调查报告中合理运用图表可以有效地表达数据资料。成功运用图表的关键在于清晰、简洁地表述报告所要传达的信息。图表的选择应适合数据的表述目的。一般图表包括：表格（多种答案的列表分析、选择题列表分析、交叉列表分析、横列表分析）、饼形图、柱状图、流程图和照片等。



本章小结

统计调查是统计整理和统计分析的基础，是保证统计工作质量的基本环节，要求统计调查做到准确、及时和全面。在统计调查之前要制定一个科学、合理的统计调查方案。

统计调查报告是对所研究的社会经济现象进行大量观察、深入研究以后，反映现象的是非曲直及其本质规律的一种专业应用文，是统计调查工作的最终成果，是对社会经济现象从感性认识到理性认识飞跃过程的反映。



复习思考题

1．专门调查有哪几种组织形式？试举例说明。
2．简述统计调查的概念及其基本原则。
3．统计调查的时点和时限有什么区别？
4．抽样调查有什么特点？请说明抽样调查与普查的关系。
5．重点调查中重点单位的含义是什么？
6．统计调查报告的种类有哪些？



实训基础训练

一、单项选择题

1．对某公司 1000 名职工的收入情况进行调查，从中随机抽取了 100 名职工发放了调查问卷，这种调查属于（　　）。

A．普查　　B．重点调查

C．典型调查　　D．抽样调查

2．对一批商品进行质量检验，最适宜采用的方法是（　　）。

A．全面调查　　B．抽样调查

C．典型调查　　D．重点调查

3.2000 年 11 月 1 日零时，我国统计部门对全国人口情况进行调查，这种调查属于（　　）。

A．普查　　B．重点调查

C．典型调查　　D．抽样调查

4．下列调查中，调查单位与填报单位一致的是（　　）。

A．企业设备调查　　B．人口普查

C．农村耕地调查　　D．工业企业现状调查

5．（　　）属于全面调查。

A．普查　　B．重点调查

C．典型调查　　D．抽样调查

6．统计调查时间是（　　）。

A．调查工作的时限　　B．调查资料所属时间

C．调查登记的时间　　D．调查期限

二、判断题

1．我国人口普查每 10 年进行一次，它属于经常性调查。（　　）

2．全面调查和非全面调查是根据调查结果所得到的资料是否全面来划分的。（　　）

3．统计报表制度是国家规定的一种统计报告制度。（　　）

4．我国人口普查的总体单位和调查单位都是每一个人，而填报单位是每户。（　　）

5．调查时间是指调查工作的起止时间。（　　）

6．采用重点调查法搜集资料时，选择的调查单位是标志值较大的单位。（　　）



课外延伸阅读

民间统计调查

一、民间统计调查的含义

1．民间统计调查是非政府统计调查

民间统计调查是由企业事业单位、社会团体或非法人单位及个人以自己的名义或接受委托进行的统计调查。也就是说，民间统计调查是由不具有行政管理职能的其他组织和个人组织进行的统计调查。民间统计调查的规模一般比较小，调查结果往往是微观的，一般是为企事业单位、社会团体生产经营服务或科学研究提供依据的，大多数民间统计调查是以营利为目的的。民间统计调查项目是由调查者或委托调查者根据自己的需要确定的，民间统计调查所使用的调查表是由民间统计调查的调查者或委托者设计制作的，调查经费也是由民间统计调查的调查者或委托者自己负担的。

2．民间统计调查采用的方法是统计方法

利用统计方法进行社会调查分为统计调查、统计资料整理、统计分析三个阶段。民间统计调查涵盖以上利用统计方法进行社会调查的全过程。无论三个阶段都进行，还是只实施其中的两个阶段或一个阶段，都属于民间统计调查。例如，某机构只实施统计调查，而不对资料进行整理和分析，如某市场调查公司将其利用统计方法搜集的原始资料提供给委托方。再如，某机构只做统计资料整理工作，如某统计师事务所接受企业委托，对现有的政府统计资料或已有的民间统计资料进行再加工；某机构利用历史和现有的统计资料，对某一经济社会现象做出预测，并公布其结果。

3．民间统计调查是社会调查

民间统计调查是社会调查涵盖了两层意思。

一是民间统计调查是对社会现象进行的调查。民间利用统计方法进行调查有两类：一类是对自然现象进行的调查，如对气温、植物生长等进行的统计调查；另一类是对社会现象进行的调查。民间统计调查是专指对社会现象进行的调查，包括社会、经济调查两部分。社会部分一般包括对人类社会的政治、法律、道德、文化、体育、卫生等现象进行的统计调查；经济部分一般包括对人类社会的生产、分配、流通、消费等现象进行的统计调查。我国现今的民间统计调查大部分是对经济和市场状况的调查，对社会部分的调查项目还比较少。但随着我国社会政治经济的发展和改革开放的不断扩大、随着人们对政治生活环境要求的不断提高，对社会部分的调查项目会越来越多。

二是民间统计调查是面向社会进行的调查。也就是说，民间统计调查的对象或服务的主体与进行民间统计调查活动的机构之间的关系是平等的民事主体关系。例如，某调查公司向其他单位或本单位以外的人员发送调查表，某统计师事务所接受另一单位的委托进行统计咨询服务。如果不是对平等的民事权利主体进行的调查，就不是民间统计调查。例如，企业事业单位内部的统计调查，社会团体对其会员单位按照社团章程进行的统计调查，都不是民间

统计调查。

二、民间统计调查的特征

1．自愿性

政府统计调查是一种强制性的调查，它是由国家授权进行的，公民和法人有接受调查的义务。民间统计调查是民事行为，民间统计调查机构作为民事法律关系的主体，与其他民事法律关系主体（包括被调查者、委托调查者、资料所有者、资料使用者等）产生的民事法律关系基于民事法律的自愿、平等原则，因此民间统计调查是自愿性调查。

民间统计调查的自愿性包括以下两方面的含义。

（1）对于被调查者而言，是否接受民间统计调查是自愿性的、非义务性的，是否对民间统计调查的调查表进行填报也是自愿性的、非义务性的。这就要求民间统计调查机构在对被调查者进行调查时，不得进行强制性调查，不得隐瞒调查目的，不得对被调查者有所损害。

（2）对于调查者而言，只要不违反国家的法律、法规和有关的规章，就可以根据调查者或委托者的意愿进行调查，民间统计调查结果的所有者可以根据自己的意愿使用、转让和买卖民间统计调查结果。这是因为民间统计调查行为本身属于民事行为，在接受国家统计行政管理部门管理的同时，还要被民事法律所调整。但是，向社会公布民间统计调查结果则必须在此之前经政府统计部门审批，不得随意进行。

2．时效性

由于民间统计调查是建立在市场经济体制下的，一般是以追求利益为目的的，所以民间统计调查较之于政府统计调查，时效性更为明显和突出，主要表现在两方面。一是快速、灵活。与政府统计调查一样，及时性也是民间统计调查的基本原则。民间统计调查较之政府统计调查来说，一般反映的是社会经济的具体运行状况，主要是为微观经济主体和社会研究主体服务的。社会经济的具体运行状况一般变化较快、形式各异，而服务的主体要求也比较急迫，因此决定了民间统计调查必须快速、灵活。同时，民间统计调查项目一般规模较小、内容单一、操作简便，这有利于民间统计调查机构能及时、灵活地组织和调整统计调查项目，快速地反映社会政治经济运行情况。二是追求效益。民间统计调查行为属于民事行为，一般是以营利为目的的。因此，在民间统计调查活动中要讲求经济效益，以最小的投入获得最大的收益。

三、民间统计调查的作用

1．民间统计调查弥补了政府统计调查的不足

政府统计是对整个国民经济和社会发展情况进行统计调查、统计分析，提供统计资料和统计咨询意见，实行统计监督。因此，国家在确定进行哪些统计调查时主要从国家宏观决策和管理的需要出发，反映的是国情国力和社会经济发展的宏观运行情况，政府统计结果的生产周期比较长。政府统计对经济发展运行的大量微观信息并不涉及，也不可能涉及。

我国实行改革开放以来，随着社会进步和经济的发展，人们对社会经济发展运行的微观信息的要求越来越多、越来越高，主要表现在以下几个方面。一是随着我国社会主义市场经济体制的建立，经济运行已由卖方市场转向买方市场，其结果是，一方面，老百姓迫切需要了解与自己生活密切相关的经济信息和产品信息；另一方面，作为生产企业和商家，为自身

的发展和竞争考虑，也需要了解与自己生产经营有关的信息。二是随着社会的进步，人们的生活观念也发生了很大的变化，人们迫切需要了解自己的生活环境。三是随着人们生活水平的不断提高，人们越来越需要了解政治生活环境和文化生活环境。在这些方面，民间统计调查发挥了积极的、不可替代的作用。民间统计调查以其快速、灵活的方法，为社会提供大量的统计信息。这些统计信息是政府统计所涉及不到或政府统计因统计力量的不足而暂时顾及不到的，因此民间统计调查起到了拾遗补缺的作用。

2．民间统计调查普及和提高了公民、社会的统计知识和统计意识

民间统计调查项目一般是微观的，统计调查结果与人们的政治、经济、文化生活关系更加密切，相对于政府统计调查而言，更直接地面对某一具体的公民或单位，调查项目往往也是公民或单位感兴趣的，公民或单位的参与程度更加广泛。另外，民间统计调查是在完全公开、平等、竞争的环境下进行的。

政府统计调查是政府统计机构行使的统计行政管理职能，是在一个相对封闭的系统中进行的，政府统计调查除每十年一次的人口普查中 1%人口抽样调查由于调查对象的特殊性而由公民直接参与以外，几乎所有的统计报表都是由国家机关、企事业单位、社会团体按一定的行政管理职责，由本单位内部负责统计工作的人员报送。因而，政府统计所涉及的人员范围是狭小的，人们对统计的了解是十分有限的。

通过民间统计调查活动的进行，使人们直接与统计调查相接触，直接成为被调查者，直接感受和体会统计调查过程，并直接接受和使用统计调查结果，对统计有了更加深刻的认识，所以民间统计调查的社会影响力是十分广泛的，能够起到“普及和提高公民、社会的统计知识和统计意识的作用”。

资料来源：百度百科

http://baike.baidu.com/link?url=PywxI3zaT2zYPK_7ZIdNI471pyNEv9j_dZnyW9CbyVrtQiwT9a5jeCOrBzX4jrUiDLymDilKrvMu4zbio5Ceyq



第3章 统计整理

学习目标要求

1. 了解统计整理的重要意义、表现形式和整理步骤，掌握统计分组的技术运用。

2. 理解分布数列的概念和种类，熟悉变量数列的编制方法；掌握统计表与统计图的构成、种类及绘制要求。

重点难点提示

1. 统计分组的技术方法、变量数列的编制。

2. 统计表的编制规则和统计图的绘制要求。

教学引导案例

投资增速回落但结构持续优化

2016 年 1～7 月，全国完成固定资产投资（不含农户）311 694 亿元，增长 8.1%，增速比上半年回落 0.9%。尽管投资增速回落，但投资结构持续优化。

一、服务业投资保持较快增长，科教文卫投入明显加大

随着我国产业升级的不断推进和优化，目前服务业投资已经成为促进投资增长的第一大推动力，在稳增长中发挥着重要作用。1～7 月，第三产业投资 179 362 亿元，增长 10.8%，增速比第二产业高 7.3%，所占比重为 57.5%，比第二产业高 18%。特别是在高技术服务业、科教文卫等领域，投资均保持良好的增长态势，1～7 月高技术服务业投资 7200 亿元，增长 18.6%，增速比上半年高 2.2%；在科教文卫领域投资 14 680 亿元，增长 17.3%，其中，卫生和社会工作、教育投资分别增长 23.6%和 23.2%。

二、工业投资结构优化，转型升级稳步推进

（1）工业高技术产业投资增速加快。1～7 月，工业高技术产业投资 11 809 亿元，增长 11.7%，增速比上半年提高 0.5%，比全部工业投资高 8.3%；占全部工业投资的比重为 9.8%，比去年同期提高了 0.7%。

（2）工业技改投资增长较快。1～7 月，工业技改投资 48 091 亿元，增长 13.8%，增速比全部工业投资高 10.4%；占全部工业投资的比重为 39.9%，比去年同期提高了 3.7%。技术进

步已成为促进工业投资增长的重要动力源泉。

（3）面向市场需求的工业消费品投资稳步增长。1～7 月，消费品制造业投资 28 107 亿元，增长 8%，增速比全部工业投资高 4.6%。其中与居民吃穿用等日常消费相关的行业投资保持了良好的增长态势，如文教、工美、体育和娱乐用品制造业投资增长了 14.3%，食品制造业投资增长了 14.2%，纺织业投资增长了 11.7%，医药制造业投资增长了 11.3%。

（4）高耗能行业投资增速持续回落。1～7 月，采矿业投资下降 22.9%，降幅比上半年扩大 3.2%。六大高耗能行业投资增长 4.6%，增速比上半年回落 0.7%，其中，化学原料及化学制品制造业投资下降 5.5%，非金属矿物制品业投资下降 1.5%，有色金属冶炼及压延加工业投资下降 5.1%。

三、“补短板”领域投资较快增长，薄弱环节投资继续加强

（1）农田水利投资快速增长。1～7 月，农业投资增长 32%，增速比上半年提高 0.5%，比全部投资高 23.9%；占全部投资的比重为 1.6%，比去年同期提高 0.3%。水利管理业投资增长 23.9%，增速比全部投资高 15.8%；占全部投资的比重为 1.5%，比去年同期提高 0.2%。

（2）信息传输、交通运输业投资保持良好的增长态势。1～7 月，信息传输业投资增长 35.4%，增速比上半年提高 0.2%，新一代信息技术的快速发展正在改变相关行业的供给格局，有效提升了服务效益。同样与传输有关，1～7 月交通运输领域投资也保持了良好的增长态势，增长 11.7%，增速比全部投资高 3.6%。

（3）生态保护和环境治理业投资高速增长。1～7 月，生态保护和环境治理业投资增长 32.4%，增速比上半年提高 5.2%，比全部投资高 24.3%。绿色发展理念正在成为投资新焦点。

（4）公共设施管理业投资稳步推进。1～7 月，公共设施管理投资增长 24.8%，占全部投资的比重为 9.1%，比去年同期提高 1.2%；对投资增长的贡献率为 24.2%，拉动投资增长 2%。

思考：

1. 该案例是对 2016 年 1～7 月的投资数据进行科学的统计整理以后，得出的统计分析结果。这些数据给我们带来哪些警示？

2. 统计整理的好坏能够反映统计工作的实质内容，决定着能否为今后的决策提供科学可靠的依据。统计整理应该如何开展？

资料来源：中华人民共和国统计局
http://www.stats.gov.cn/tjsj/sjjd/201608/t20160812_1387886.html

第一节　统计整理概述

在统计调查中，通过运用一定的统计调查方法，人们取得了大量能够说明总体中个体特征的原始资料。但是这些原始数据资料是分散、不系统的，只能显示各个调查单位的具体情况，不能说明被研究总体的全貌，因此，必须对这些资料进行加工整理才能认识事物的总体特征和内部联系。

为了使人们正确认识社会经济现象总体数量特征，必须按照科学的原则对这些个别的、分散的资料运用科学的方法进行加工整理，使之系统化和条理化，才能认识事物的总体特征

及内部联系，以便对总体做出概括性的说明。使统计调查资料系统化、条理化的过程就是统计资料的整理工作。

统计资料整理是统计工作的第三阶段，它在整个统计工作过程中起着承前启后的作用，它既是统计调查的继续和深化，又是统计分析的基础和前提，是统计调查和统计分析的连接点。

一、统计整理的含义

根据统计研究任务的要求，对统计调查所搜集到的原始资料进行科学的分类、汇总，使其条理化、系统化的工作过程就是统计整理。对于已经整理加工过的资料（次级资料）进行的再整理也属于统计整理。

统计调查可以获得大量调查单位的原始资料，这些资料仅说明各个单位的具体情况，是分散的、不系统的，只能反映个别事物的某一个侧面，不能揭示总体事物的本质特征，也不能从量的方面反映事物发展变化的规律性，因此，需要对统计调查获得的原始资料进行加工、分类和汇总，使其条理化、系统化。资料整理是资料研究的重要基础，是提高调查资料质量和使用价值的必要步骤，是保存资料的客观要求。

统计整理的任务就是根据统计研究的目的和要求，借助综合指标，有组织、有计划地对统计调查中搜集到的资料进行加工处理，使其成为系统化、条理化的综合资料，对总体内部规律性、相互联系、结构关系做出概括的说明。统计整理使采集到的分散、不系统的原始数据成为有序的统计信息。

例如，人口普查中搜集到的人口资料，只能说明每一个人的具体情况，如每个人的姓名、性别、年龄、文化程度等。必须对人口总体中每个人的资料进行整理、分组、汇总等加工处理，才能得到人口总体的综合情况，从而了解人口总体的规模、结构、增减变动状况等信息，达到对人口总体全面、系统的认识。

对统计资料进行整理是人们对社会经济现象从感性认识上升到理性认识的过渡阶段，是统计工作中的一个十分重要的工作阶段，在整个统计工作过程中起着承前启后的作用，它既是统计调查的继续和深入，又是统计分析的基础和前提。换一句话说，统计整理的正确与否、质量好坏，将直接影响统计对社会经济现象数量描述的准确性和数量分析的真实性。

不恰当的加工整理往往使调查得来的丰富、准确、全面的资料失去应有的价值，从而歪曲事情的真相，使人们得出错误的结论。因此，采用科学的方法进行统计整理是顺利完成统计分析任务的前提。统计整理是实现由对个别现象的认识过渡到对总体现象的认识、由对事物表象的认识过渡到对其本质与内在联系的全面深刻认识，由感性认识上升到理性认识的过程，是达到统计研究目的的重要环节。

二、统计整理的方法

统计整理的方法主要有统计分组、统计汇总和编制统计表。

1. 统计分组

统计整理的第一步是根据研究任务的要求，确定要对调查所得的原始资料进行哪些分组或分类，在分组的基础上确定应该汇总得到哪些统计指标。

2. 统计汇总

统计汇总即在统计分组的基础上，把总体单位各种标志的标志值汇总起来。汇总主要分为手工汇总和电子计算机汇总。

3. 编制统计表

编制统计表即把汇总的资料按一定的规则在表格上表现出来。

三、统计整理的内容和步骤

（一）统计整理的内容

进行统计整理的时候主要完成以下三方面的工作：

首先，根据研究任务的要求，选择应整理的指标，并根据分析的需要确定具体分组；

其次，对统计资料进行汇总、计算；

最后，通过统计表描述汇总的结果。

在统计整理中，抓住最基本的、最能说明问题本质特征的统计分组和统计指标对统计资料进行加工整理，是进行统计整理必须遵循的原则。

（二）统计整理的步骤

统计整理的步骤由内容来决定，大体分为以下几个步骤。

1. 设计整理方案

整理方案与调查方案应紧密衔接。整理方案中的指标体系与调查项目要一致，或者是其中的一部分，绝不能矛盾、脱节或超越调查项目的范围。整理方案是否科学，对于统计整理乃至统计分析的质量都是至关重要的。

2. 对调查资料进行审核

在汇总前，要对调查得来的原始资料进行审核，审核它们是否准确、及时、完整，发现问题，加以纠正。统计资料的审核也包括对整理后次级资料的审核。如果发现调查资料在及时性或完整性方面存在问题，要及时督促填报单位按时、及时报送；有遗漏、填写不全的项目时要及时通知填报单位补报、补填。



小贴士

统计资料的审核

为了保证统计资料整理的质量，在对统计资料进行整理前，先要对所有的原始资料进行严格的审查核对。如果发现问题，要及时更正，这是统计整理工作过程中十分重要的一环。审核的内容包括资料的准确性、完整性和及时性三方面。

一、准确性审核

准确性审核是统计审核的重点。准确性审核有两种审核方法。

1. 计算检查

计算检查就是利用平衡或加总关系审核调查表或报表中各项数字在计算方法和结果上有无差错。例如，各横行纵栏的合计是否错误、数字的计量单位是否正确、各指标的计算方法是否恰当等。

2. 逻辑检查

逻辑检查就是从调查项目的相互关系中、从被调查事物的特点中、从现象发展变化的常态中审核调查资料的内容是否合理、有无相互矛盾和不符合实际的地方。

二、完整性审核

完整性审核就是审核所有被调查单位的资料是否齐全、是否按规定的项目和份数上报；所搜集的资料是否达到规定的调查单位数目，调查资料中的各项目填写是否齐全。因为任何单位的资料不报、缺报都会影响整个汇总工作的进行。

三、及时性审核

及时性审核就是审核调查资料是否按规定的时间报送，并检查未及时报送的原因。

3. 进行科学的统计分组

用一定的组织形式和方法对原始资料进行科学的分组，是统计整理的前提和基础。按分组的要求，对各项数字进行汇总，计算分组单位数、总体单位数、分组标志总量和总体标志总量。在统计整理过程中，对大量的原始资料进行分组、汇总和计算是一项主要的工作。

4. 对汇总后的调查资料进行审核

对分组后的资料进行汇总和必要的计算，就使得反映总体单位特征的资料转化为反映总体数量特征的资料，然后对汇总后的资料进行审核，主要从以下几方面进行审核。

①复计审核：对每个指标数值进行复核计算。

②表表审核：审核不同统计表上重复出现的同一指标数值是否一致；对于统计表中互有联系的各个指标数值，则审核它们之间是否衔接和符合逻辑性。

③表实审核：将汇总得到的指标数值与了解的实际情况联系起来进行检查。

④对照审核：对某些统计、会计、业务三种核算都进行计算的指标数值，应进行相互对照检查，看数字是否相同，以便从中发现可能出现的错误。在审核过程中发现错误时，应查明原因，及时更正。

5. 编制统计表并绘制统计图

把整理好的统计资料用统计表或统计图的形式表现出来，以简明扼要地表现社会经济现象在数量方面的具体特征和相互关系。统计表是统计资料整理的结果，也是表达统计资料的重要形式之一，根据研究的目的可编制各种统计表。一般来说，统计表是统计整理的最后环节，当汇总的结果体现在统计表或统计图上时，就意味着整理过程结束了。

第二节 统 计 分 组

统计整理的主要工作内容是对调查得到的资料进行分组、汇总和计算，其中统计分组是最基本的工作，是保证分类、汇总科学合理的基础。

一、统计分组的概念和作用

（一）统计分组的概念

统计分组是根据统计研究的目的和要求，按照一定的标志将总体划分为若干个不同的类别组，使组与组之间有较明显的差异，使调查来的大量无序的统计资料成为有序的、层次分明的、能表现总体属性和特征的资料。

统计分组是在统计总体内部进行的一种特定分类，它同时具有两方面的含义：对总体而言是“分”，即将总体分为性质相异的若干部分；对个体而言是“合”，即将在某些方面性质相同的个体组合起来。

统计总体中的各单位，一方面，在某一个或几个标志上具有相同的性质，可以被结合在同一性质的总体中；另一方面，又在其他标志上具有彼此相异的性质，从而又可以被区分为性质不同的若干个组成部分。

（二）统计分组的作用

统计分组是统计整理的主要方法，统计分组的作用主要体现在以下三个方面。

1. 划分现象的类型

统计分组的根本作用在于区分现象的质。

任何一批数据都存在差异，在进行统计分组之前这种差异处于无序状态，显现不出来，通过统计分组，反映出了统计总体的基本性质和特征。分组实际上就是按差异的大小进行分类，差异小的归入一组，差异大的归入其他的组。因此，统计分组的结果使组内的差异缩小、组与组之间的差异扩大。

统计分组的过程就是区别事物性质的过程。要了解各种社会经济现象的性质、特点及其相互关系，必须按某种标志把它们划分为性质不同的部分。

2. 反映现象的内部结构

统计往往对总体按某一标志进行分组，并计算总体内各组成部分占全体的比重，以说明各个组成部分在总体中的分布状况，反映现象的内部结构和结构变化，从而揭示现象的性质和发展变化的规律。这种分组既可以采用品质标志分组，又可以采用数量标志分组。

3. 分析现象之间的依存关系

社会经济现象不是孤立存在的，各现象之间存在广泛的联系和制约关系，一种现象的变化常是另一种现象变化的原因或结果。通过统计分组，可以揭示现象之间的依存关系。例如，居民收入和消费支出之间的关系、商品销售额与流通费用率之间的关系等。这些方面的依存关系都可以利用统计分组划分事务类型，说明影响因素对结果因素的作用程度。

统计分组是基本的统计方法之一，在统计资料的整理和分析中都要广泛应用分组。而分组的好坏直接关系到统计整理质量，关系到统计分析的结论是否正确。从某种意义上讲，没有科学的统计资料整理，就没有科学的统计分析。

二、分组标志的选择和各组界限的划分

统计分组中关键的问题在于选择分组标志和划分的各组界限，而选择分组标志是统计分组的核心问题。

（一）选择分组标志

分组标志是统计分组时所依据的标志。正确选择分组标志必须从对客观事物的定性认识出发，力求选择与统计研究目的和任务有密切关系并能够反映事物本质特征的标志作为分组标志。

分组标志就是将统计总体区分为各个性质不同的组的标准或根据，任何一种社会现象客观上都有许多不同的标志。对同一总体的资料根据不同的标志进行分组时，会产生不同的结论，因此，为确保分组后的各组能够正确反映事物内部的规律性，选择分组标志时应遵循以下原则。

1. 根据统计研究的目的与任务选择分组标志

在对社会经济现象进行研究时，可以根据不同的研究目的或任务从不同的角度进行研究，相应地要选择不同的分组标志进行分组。例如，以全国总人口为总体进行研究时，这个研究对象就有很多标志，如性别、居住地、年龄、学历等。

在具体研究过程中到底应该采用哪种标志进行分组，就要看研究的目的。如果研究人口的性别构成状况，则按照性别进行分组；如果要研究人口的文化程度，就要按照学历进行分组了。

2. 根据被研究现象的本质特征选择分组标志

同一研究总体，由于研究的目的和任务不同，所选择的分组标志也不同。例如，以某物

流企业的员工为总体，如果研究的目的是分析员工的文化素质，就应选用员工的文化程度作为分组标志，这样才能分析出具有大学、中学、小学等不同文化程度的员工各占多大比例；如果研究的目的是分析员工的劳动力素质，则应以员工的年龄作为分组标志，这样才能分析出员工中老年、中年、青年的比例。

3. 根据现象所处的历史或经济条件选择分组标志

社会经济现象的特征有多种，选择分组标志时需要根据被研究现象的特点选择主要的、能反映事物本质特征的标志作为分组标志。例如，物流企业规模可以用许多标志反映，如职工人数、产品产量、产值、生产能力、经营资金量、利税额等，但在实际统计工作中，一般使用生产能力、经营资金等作为分组标志。

另外，在选择分组标志时，还要遵循互斥性和穷尽性两个原则。互斥性原则是指统计分组必须保证总体的每个单位只能属于其中的一个组，不能出现重复统计的现象。穷尽性原则是指统计分组必须保证总体的每个单位都能归入其中的一个组，各个组的单位数之和等于总体单位总数，总体的指标必须是各个单位相应标志的综合。

（二）各组界限的划分方法

根据分组标志的不同特征，统计总体可按品质标志分组，也可按数量标志分组。

1. 按品质标志分组

选择分组标志，只解决了按什么分组的问题，而一个总体应分成几个组、各组之间的界限划分在哪里为宜，便又成为关键问题。

按品质标志分组时，确定组限有两种情况。第一种情况，组限是自然形成的或比较明显的，如人口按性别、文化程度、民族等分组，由于组限自现、组数固定，因而容易确定。第二种情况，分组界限较难以确定，如人口按职业分组、企业按行业分组、产品按用途分组等，不是在每次分组时都能辨析清楚的。因此，为了保证这类分组的准确划一，这种比较复杂的属性的分组，根据国家有关部门制定的各种分类标准与分类目录分组就方便多了。

在实际工作中，对于一些比较复杂的分组，往往根据研究任务的要求对经济现象进行具体的、深入的了解，国家统计局及中央有关部门经研究统一制定了各种分类目录与分类标准，如《工业部门分类目录》、《国民经济行业分类和代码》、《工业产品目录》、《经济类型划分规定》等，供全国各地区、各部门、各单位分类时使用，以保证各种分类的统一性和完整性，见表 3-1。

表 3-1　2015 年上海市全社会固定资产投资总额按产业类型分组表

指　标	绝对值 （亿元）	比上年增长 （%）
全社会固定资产投资总额	6352.70	5.6
第一产业	3.95	-66.7
第二产业	958.84	-17.1
第三产业	5389.91	11.2

2. 按数量标志分组

按数量标志进行分组就是根据统计研究的目的，选择反映事物数量差异的数量标志作为分组标志，在数量标志值的变异范围内划定各组数量界限，将总体划分为性质不同的若干个组成部分。

按数量标志分组的目的并不是单纯确定各组在数量上的差别，而是要通过数量上的变化来区分各组的不同类型和性质。

按数量标志分组的结果形成变量数列。在统计整理和统计分析中，变量数列应用得相当广泛，以此来观察某种指标的变动及分布情况。

按数量标志分组过程中，根据变量值取值范围不同，分组的形式可以分为单项式分组和组距式分组。

（1）单项式分组

单项式分组，即每一组只包含一个变量值。这种分组形式只适用于离散变量，而且要求在离散变量的变动范围较小、变量值个数较少时使用。例如，居民家庭按儿童数或人口数分组，均可采用单项式分组。一般情况下分组组数等于变量的取值个数，各组之间的界限很明确，不需要人为划分。

（2）组距式分组

组距式分组，即在变量值变异幅度较大时，将变量值取值范围人为地划分为若干个区间，变量在同一区间内取值的现象归为一组，区间的距离称为组距。这样的分组形式下每一组中包含若干个变量值，适用于所有的连续变量和取值范围较大的离散型变量。例如，企业按职工人数分组、耕地按粮食平均单产分组、商店按销售额分组、学生按学习成绩分组等都是组距式分组。

在组距式分组中，相邻组既可以有确定的上下限，也可将相邻组的组限重叠。连续变量由于不能一一列举其变量值，只能采用组距式的分组方式，且相邻的组限必须重叠。例如，以总产值、商品销售额等为标志进行分组，就只能选择相邻组限重叠的组距式分组。

在相邻组组限重叠的组距式分组中，若某单位的标志值正好等于相邻两组的上下限的数值，则一般把此值归并到作为下限的那一组（适用于连续变量和离散变量），称为“上限不在内”。

组距式分组使资料的真实性受到一定程度的损害。组距式分组的假定条件是：变量在各组内的分布都是均匀的（各组标志值呈线性变化）。通过组距式分组以后，把各组内部各单位的次要差异抽象掉，而把各组之间的主要差异突出来，这样，各组分配的规律性可以更容易地显示出来。

根据这个道理，如果组距太小、分组过细，则容易将属于同类的单位划分到不同的组，因而显示不出现象类型的特点；但如果组距太大、组数太少，则会把不同性质的单位归并到同一组中，失去区分事物的界限，达不到正确反映客观事实的目的。因此，组距的大小、组数应根据研究对象的经济内容和标志值的分散程度等因素确定，不可强求一致。

例如，对于学生某一门课程的考试分数，一般的分组是这样的：60 分以下是一组，60～70 分一组，70～80 分一组，80～90 分一组，90 分以上一组。60 分包括 60 分以上就是我们

平常所说的及格分，60 分以下就是不及格，这样分组能够反映客观事物的本质，见表 3-2。

表 3-2　某班学生英语考试成绩分组表

成绩（分）	人数（人）
60 以下	6
60～70	9
70～80	17
80～90	10
90～100	8
合计	50

进行组距式分组时，根据各组间组距是否一致，组距式分组可以细分为等距分组和不等距分组。等距分组就是各组保持相等的组距，也就是说各组标志值的变动都限于相同的范围；不等距分组即各组组距不相等的分组。

统计分组时采用等距分组还是不等距分组取决于研究对象的性质特点。在标志值变动比较均匀的情况下宜采用等距分组。等距分组便于各组单位数和标志值的直接比较，也便于计算各项综合指标。在标志值变动很不均匀的情况下宜采用不等距分组。不等距分组有时更能说明现象的本质特征。

三、统计分组体系

统计分组体系，就是根据统计分析的要求，通过对同一总体进行不同分组，形成一系列相互联系、相互补充从多方面反映总体内部关系的体系。

分组体系有平行分组体系和复合分组体系两种。平行分组体系是选择两个或两个以上的标志对总体分别进行简单分组后所形成的体系；复合分组体系就是复合分组后形成的体系。

1. 简单分组和平行分组体系

总体只按一个标志分组称为简单分组。例如，我国人口总体按照性别分组就是简单分组。对同一个总体选择两个或两个以上的标志分别进行简单分组就形成平行分组体系。

例如，为了解我国人口总体的基本特征，我们将人口总体按性别、民族、文化程度进行分组，形成的平行分组体系如下：

（1）按性别分组　　（2）按民族分组　　（3）按文化程度分组

男　　汉族　　文盲和半文盲

女　　回族　　小学毕业

苗族　　中学毕业

壮族　　大学毕业

……　　大学以上

平行分组体系的特点：每个分组固定一个分组标志，即只考虑一个因素的差异对总体内部分布情况的影响，同时各个简单分组之间彼此独立，没有主次之分，互相不影响。

2. 复合分组和复合分组体系

对同一总体选择两个或两个以上分组标志层叠起来进行分组，即先按一个主要标志分组，再按另一个从属标志在已分好的各组中再次分组，叫作复合分组。复合分组所形成的分组体系叫作复合分组体系。例如，为了认识我国高等院校在校学生的基本状况，可以同时选择学科、学制、性别三个标志进行复合分组，得到如下的复合分组体系：

理科学生组	文科学生组
本科学生组	本科学生组
男生组	男生组
女生组	女生组
专科学生组	专科学生组
男生组	男生组
女生组	女生组

复合分组体系的特点：每一次分组除了要固定本次分组标志对分组结果的影响外，还要固定前一次或前几次分组标志对分组结果的影响；各个分组标志之间有主次之分。复合分组体系可以从不同角度了解总体内部的差别和关系，因此比平行分组体系更能全面、深入地分析和研究问题。

对于复合分组体系，复合分组的组数等于各简单分组组数的连乘积，如果复合分组的选择标志过多，就会使复合分组体系过于庞大，将增加分组的难度，更不容易反映现象的本质特征，制表也不方便。所以，复合分组时分组标志不宜过多。

第三节　分 配 数 列

分配数列是进行统计分组的必然产物，是统计整理结果的一种重要表现形式，也是统计描述和统计分析的重要内容。它可以表明总体的分布特征和内部结构，并为研究总体中某种标志的平均水平及其变动规律提供依据。

一、分配数列的概念和构成要素

（一）分配数列的概念

分配数列是在统计分组的基础上，将总体的所有单位按组归类整理，并按一定顺序排列而形成的总体中各个单位在各组间的分布，又称分布数列或次数分配。

例如，某高校 300 名教职工按文化程度分组形成的分配数列如表 3-3 所示。

表 3-3　某高校教职工文化构成情况

按文化程度分组	教工人数（频数）	占总数的比率（频率）（%）
博士及以上	45	15.0
硕士	175	58.3
大本	45	15.0
大本以下	35	11.7
合计	300	100

（二）分配数列的构成要素

1. 构成要素

分配数列包括品质分配数列和变量分配数列，无论是哪种分配数列，都是由两个基本要素构成的：分组的组别（也叫作各组的名称）和各组的次数或频率。

分配数列中分布在各组中的个体单位数叫作次数，又称频数；各组次数（即各组单位数）占总次数（即总体单位数）的比例叫作比率或频率，各组次数之和等于总次数，各组频率之和等于 1 或 100%。

两种分配数列构成要素的不同之处仅有一点，即品质数列组的名称使用文字表示标志的属性差异；而变量分配数列组的名称则是用标志值（即变量的不同水平）表示的数量变异界限，见表 3-2 和表 3-3。

2. 要素的意义和作用

分配数列中各组的名称是表明标志变异范围及其变异程度界限的。而次数（频数）和频率，前者是以绝对数的形式表现各组的总体单位数目，后者则是以相对数形式表现的总体单位数目。在变量数列中，次数越大的组的标志值对总体指标计算所起的作用越大；反之，次数越小的组的标志值所起的作用也越小。

频率与次数所起的作用根本上是相同的，不同的是它还可以表明各组标志值对总体的相对作用程度。这种相对作用程度的具体数值也是各组标志值在总体中出现的概率。为研究整个变量数列的次数分配状况和进行某种统计计算，统计工作中还常计算累计次数及其频率分布。将变量数列中各组的次数和频率逐组累计相加而形成累计次数分布，它表明总体在某一变量值的某一水平上下总共包含的总体次数和频率。

一般来说，累计次数和累计频率的计算方法有两种：向下累计和向上累计。向下累计又称较大制累计，将各组次数或频率由变量值高的组向变量值低的组累计，各累计数的意义是各组下限以上的累计次数或累计频率；向上累计又称较小制累计，是将各组的次数或频率由变量值低的组向变量值高的组累计，各累计数的意义是各组上限以下的累计次数或累计频率，如对某班 50 名学生统计学考试成绩分组，见表 3-4。

表 3-4　某班 50 名学生统计学考试成绩分布表

按考试成绩分组	人　数		向上累计		向下累计	
	频数（人）	频率（%）	频数（人）	频率（%）	频数（人）	频率（%）
60 以下	2	4	2	4	50	100
60～70	6	12	8	16	48	96
70～80	12	24	20	40	42	84
80～90	20	40	40	80	30	60
90～100	10	20	50	100	10	20
合计	50	100	—	—	—	—

二、分配数列的种类

根据分组标志的不同，分配数列分为品质分配数列和变量分配数列两种。

1. 品质分配数列

按品质标志分组形成的分配数列叫作品质分配数列，简称品质数列，也叫属性分布数列。例如，表 3-1、表 3-2、表 3-4 都是品质分配数列。编制品质分配数列，只要分组标志选择得好、分组标准定得恰当，则事务性质的差异就表现得比较明确，总体中各组的划分也比较容易。因而品质分配数列一般比较稳定，能准确地反映总体的分布特征。

2. 变量分配数列

按数量标志分组所编制的分配数列叫变量分配数列，简称变量数列。例如，表 3-3 就是数量分配数列。

变量数列按变量的表示方法和分组方法不同，可以分为单项式变量数列和组距式变量数列两种。

（1）单项式变量数列

单项式变量数列即指将每一变量值列为一组形成的数列，也称为单项数列，即按单项式分组所编制的变量数列。适用于变量值个数较少、变动范围较小的离散型变量。变量值数目较少并可一一列举，因此可编制单项数列，如表 3-5 所示。

表 3-5　某学生社团学生年龄分布情况

按年龄分组	人数（频数）	比率（频率）（%）
18	1	5.0
19	3	15
20	6	30.0
21	4	35.0
22	2	10.0
23	1	5.0
合计	20	100

（2）组距式变量数列

组距式变量数列是以标志值变动的一定范围作为一组的分组，即组距式分组所形成的变量数列，也称为组距数列。组距数列中的每个组不是用一个具体的变量值表示的，而是用变量值的一定变化范围（即各组变量值变动的区间）来表示。

组距数列一般适用于连续型变量及变量值变动范围较大的离散型变量。在离散型变量变动范围比较大、统计单位数又很多的情况下，若编制单项式数列，把每一变量值作为一组，则必然使分组的组数过多、各组次数过于分散，不能反映总体内部各部分的性质和差异，从而失去了编制分配数列的意义。至于连续型变量，由于变量值无法一一列举，更不能编制单项式数列。在这些情况下就需要编制组距式变量数列。

三、变量分配数列的编制

品质数列比较简单，如果分组标志选择得好、分组标准定得恰当，则事物性质的差异表现得比较明确，总体中各组也容易划分。在编制品质数列时，只要按规定的分组标准将总体单位按组归类整理即可。品质数列一般来说比较稳定，通常能准确反映总体的分布特征。这里只重点介绍变量分配数列的编制。

1. 单项数列的编制

在编制单项数列时，首先将调查所得资料按照数值由小到大的顺序排列；其次，确定各组的变量值和组数，一般有多少个变量值就有多少个组；最后汇总出各变量值出现的次数，编制单项数列。由于单项数列每组只有一个变量值，各组之间界限划分也非常明确，因此编制出的数列也更稳定。

2. 组距数列的编制

组距数列编制的难度比较大。编制过程中，一般首先把变量值按照由小到大的顺序排列，并确定全距；其次，确定组数和组距；在此基础上确定组限；最后，汇总出各组的次数及比重，编制组距式数列。

组距式数列编制重点是解决组数、组距、组限及组中值的几个要素的确定问题。

（1）组数与组距

正确确定组数和组距是组距式数列编制过程中的一个关键问题。在组距数列中我们用变量值变动的一定范围代表一个组，每个组的最大值为组的上限、最小值为组的下限，每个组上限和下限之间的距离即为组距。

组数是指某个变量数列划分为多少组，组数与组距是相互联系的，在同一变量数列中，组距的大小与组数的多少成反比，组数越多，组距越小；组数越少，组距越大。

确定组数与组距时，一般先要找出全部变量值的最大值和最小值的距离（即全距），以及大多数变量集中的区间，然后根据标志变量的分散程度及项目数量等因素来考虑组距和组数的问题。要使组距反映各组之间的数量界限，把性质相同的单位归入一组，而将性质不同

的单位划分为不同的组列，以保证编制的组距数列尽可能反映出总体分布的特征及其规律性。

对于一个统计总体来讲，分组的组数和组距必须恰当，不能太多也不能太少。组数太多，组距太小，分配数列显得很烦琐，不能反映总体分布特征；组数太少，组距太大，分配数列显得过于笼统，同样很难反映总体分布特征。

例如，对吉林省某农村居民家庭人均月消费性支出情况进行抽样调查，得到 40 户家庭人均月消费性支出（单位：元）资料如下：

304　349　355　361　376　385　388　393　397　400
400　406　409　413　414　414　416　417　422　429
432　435　441　442　450　456　464　470　472　478
482　485　488　490　504　524　535　544　566　593

这里共有 40 个变量值，最小值是 304 元，最大值是 593 元，标志变动的范围为 304～593 元，全距为 289 元。经过综合考虑，我们可以将其分成 6 个组距相等的组，则组距为 48.2，实际取组距 50 元，则可以编制组距数列，如表 3-6 所示。

表 3-6　某农村居民家庭人均月消费性支出分配数列

人均月消费性支出（元）	户　数	
	绝对数（户）	相对数（%）
300～350	2	5.0
350～400	7	17.5
400～450	15	37.5
450～500	10	25.0
500～550	4	10.0
550～600	2	5.0
合计	40	100.0

表 3-6 中的分组能较好地反映某农村居民家庭人均月消费性支出的分布状态，如果组数过少和过多，会造成组距过大或过小，都不利于显示总体内部各组的分布特征。

由此可见，编制组距式数列时，不仅要考虑各组的划分是否能区分总体内各组成部分的性质差异，还要适当地确定组距和组数，才能准确而清晰地反映总体的分布特征。在实际统计工作中确定组数时，可以根据对事物的定性分析加以确定。这需要对统计总体进行认真的分析，才能做出准确的判断。

确定组数时，还可以利用美国学者史德杰斯提出的经验公式。其公式如下：

$$组数=1+3.322\lg N$$（其中 N 为标志值项数）

不过这一公式不是在任何情况下都必须遵守的法则，它只适用于总体趋于正态分布的条件下进行等距分组的情况。

（2）组限与组中值

组距两端的数值称为组限，其中每组的起点数值称为下限，每组的终点数值称为上限。上限和下限的差称为组距，表示各组标志值变动的范围。在组距式分组中常有最小组无下限

和最大组无上限的情况，这样的组叫开口组，其中只有上限而无下限的组称为下开口组，有下限而无上限的组称为上开口组，如表 3-7 所示。

表 3-7　某月某车间工人按生产定额完成程度分组表

按生产定额完成程度分组（%）	工人数（人）	比率（频率）（%）
90 以下	5	4.8
90～100	14	13.3
100～110	37	35.2
110～120	29	27.6
120～130	15	14.3
130 以上	5	4.8
合计	105	100.0

在编制组距式分布数列之前，应对标志值的分布情况进行仔细审查，在分布较集中的标志值中确定出组距的中心位置，尽可能将总体内各单位的分布特征表现出来。

在编制组距数列时，作为各组名称的变量可以是离散变量也可以是连续变量，这两种变量组限的表示方法有所不同。

按连续变量分组划分组限时，相邻两组的组限必须重叠，如表 3-8 所示。这是由于连续变量相邻两个变量值之间可以进行无限的分割，如果上下限是两个不同的数值，那么相邻两组上下限之间就可能有很多数值无组可归，不符合穷尽性原则。因此，相邻两组上下限必须用同一个数值表示。

在统计工作中，如果遇到某单位的标志值刚好等于相邻两组上下限数值时，为避免重复计算，一般遵循“上限不在内”的原则。例如，表 3-8 中某月某车间工人按生产定额完成程度分组，其中有 90～100、100～110 两组，如果某位工人定额完成程度恰好为 100%，则应计入 100～110 这一组。

离散变量在划分组限时，相邻两组的组限可以以整数断开，如表 3-8 所示。因为离散变量不能用小数表示，相邻两个变量值可以以整数断开，因此，如果相邻两组的组限不重叠，也不会导致遗漏。

表 3-8　某地区 25 个商店按职工人数分组表

按职工人数分组	商店个数（个）	各组占总数百分比（%）
15～25	5	20.0
26～35	3	12.0
36～45	3	12.0
46～55	5	20.0
56～65	2	8.0
66～75	2	8.0
76～85	5	20.0
合计	25	100.0

组中值是上下限之间的中点数值，以代表各组标志值的一般水平。组中值的计算公式为：

组中值=（上限+下限）/2

例如，表 3-7 中，120～130 一组的组中值为（120+130）/2=125。

组中值仅存在于组距式分组数列中，单项式分组中不存在组中值。组中值的计算是有假定条件的，即假定各组标志值的变化是均匀的（与组距式分组的假定条件相同）。一般情况下，对于第一组是“多少以下”、最后一组是“多少以上”的开口组来说，组中值的计算可参照邻组的组距来决定：

缺下限开口组组中值=上限-相邻组组距/2

缺上限开口组组中值=下限+相邻组组距/2

用组中值来代表组内变量值的一般水平存在一个假定，即假定各单位变量值在本组范围内呈均匀分布，或在组中值两侧对称分布；而实际中各组变量值往往不是均匀分布的，组中值与各组的实际平均水平仍有一定的差距，它只是各组实际平均值的近似代表值。

在实际工作中，数据的分布具有不确定性。在确定组距时，可以根据标志值变异状况不同将各组的组距确定为相等的或不等的，相应地就形成了等距数列和不等距数列，也称异距数列。

等距数列是各组组距都相等的数列，一般在社会经济现象性质差异的变动比较均衡或标志变异比较均匀的条件下采用。在等距分组时，一般先确定全距、组数，然后用全距除以组数得出组距，并据此划分各组的界限。全距、组数和组距的关系用公式表达如下：

$$I = R / K, \qquad R = X_{\max} - X_{\min}$$

式中：I 为组距；R 为全距；K 为组数；$X_{\max}$ 为最大变量值；$X_{\min}$ 为最小变量值。

以上计算结果只是一个参考数值，为计算方便，实际工作中一般组距取 5 或 10 的整数倍。当然也可以先确定组距，再确定组数。

异距数列是各组组距不相等的数列，也叫不等距数列，通常适用于社会经济现象数量变动不均衡且很难用等组距的办法实现区分事物不同性质的情况。例如，对人口按照年龄分组，就不适宜采用等距分组，只能采用异距分组，如表 3-9 所示。

表 3-9　某地 18 岁以下人口按年龄分组情况表

按年龄分组（岁）	人数（万人）
一岁以下	5
1～3	7
4～6	10
7～12	12
13～18	18
合计	52

有些现象中，各类型差异很大，其标志值呈等比级差变化，则组距也应按等比级数确定。例如某地 450 家商店按营业额多少分组来反映其规模类型分布特征，如表 3-10 所示。

表 3-10　某地 450 家商店按营业额分组表

年营业额（千元）	商店个数（个）
1～10	6
10～100	54
100～1000	148
1000～10 000	215
10 000～100 000	26
100 000～1 000 000	1
合计	450

第四节　统计表和统计图

一、统计表

通过统计整理得到的反映社会经济现象总体特征的综合资料是统计工作的初步成果，需要运用一定的形式将其展示出来，以便于人们分析和利用，统计表是最常见的表现形式。经过统计汇总，得到一系列统计指标的数字资料，把这些数字按一定的逻辑顺序在表格中表现出来，这种表称为统计表。统计表是表现统计资料最常用的形式，也是统计分析的重要工具。

统计表有广义上的统计表与狭义上的统计表之分，广义的统计表泛指统计工作各个阶段以纵横交叉的线条所绘制成的用来表现统计资料的表格；狭义的统计表是专门用以表现经过整理的系统化的统计资料的表格。

（一）统计表的结构

统计表的结构从形式上看，主要由总标题、横行标题、纵栏标题、数字资料四部分构成。此外，有些统计表在表下还增列补充资料、注解、附记、资料来源、填表单位、填表人员及填表日期等内容。

总标题是统计表的名称，用来概括全表统计资料的内容，一般写在表的上端中部。

横行标题是统计表横行的名称，通常用来表示各组的名称，它代表统计表所要说明的对象，一般写在表的左方。

纵栏标题是统计表纵栏的名称，通常用来表示分组标志和指标的名称，一般写在表的上方。

数字资料是统计表中的指标数值，列在各横行和纵栏的交叉处。统计表中任何一个数字的含义都由总标题、横行标题和纵栏标题共同说明。

统计表的结构如表 3-11 所示。

表 3-11　20××年某地区工业总产值（总标题）

按经济类型分组	企业单位数（万个）	工业总产值（亿元）
国有企业	1.47	224.67
集体企业	10.36	213.21
城乡个体工业	10.12	82.05
其他经济类型工业	0.81	32.06
合　　计	22.76	551.99

横行标题　　纵栏标题　　数字资料

主　　词　　宾　　词

统计表的结构从内容来看，包括主词和宾词两个部分。主词是统计表所要说明的总体，它可以是各个总体单位的名称、总体的各个组或总体单位的全部。宾词是用来说明主词的各个统计指标，包括指标名称和指标数值。

通常把主词放在表的左方，即列于横行；宾词放在表的右方，即列于纵栏，如表 3-11 所示。但这样排列有时会使统计表的表式过分狭长或过于宽短，为了编排合理、阅读方便，也可以将主词和宾词合并排列或互换位置。

（二）统计表的种类

1. 按统计表的作用不同分类

按统计表的作用不同可分为调查表、整理表和分析表。

调查表是指在统计调查阶段用于登记、收集原始统计资料的表格。

整理表也称汇总表，是在统计整理阶段用于汇总统计资料和表现汇总结果的表格。

分析表是在统计分析阶段用于对整理所得的统计资料进行定量分析的表格。这类表格往往与整理表结合在一起，成为整理表的延续。

2. 按统计表主词分组情况不同分类

按统计表主词分组情况不同，可分为简单表和分组表。

简单表是主词未经任何分组的统计表。例如，由总体单位名称组成的一览表，主词由地区、国家、城市等目录组成的区域表，主词由时间组成的年表等，如表 3-12 与表 3-13 所示。

表 3-12　　某公司×年×月购销情况

所属地区	购进额（元）	销售额（元）
华北 华中 华南		
合计		

表 3-13　某物流企业 2012—2016 年营业额

年　份	营业额（万元）
2012	
2013	
2014	
2015	
2016	

分组表又可分为简单分组表和复合分组表。

简单分组表是指主词按一个标志进行分组后形成的统计表。

复合分组表是指主词按两个或两个以上标志层叠起来进行分组后形成的统计表。例如，表 3-14 所示的统计表就是一个复合分组表。

表 3-14　某企业工人基本情况表

	工人人数（人）	月工资总额（元）
男性		
30 岁以下		
30～50 岁		
50 岁以上		
女性		
30 岁以下		
30～50 岁		
50 岁以上		
合计		

3. 栏数编号的设计

如果统计表栏数较多，通常要加以编号。习惯上主词栏部分分别编以“甲、乙、丙、丁……”为序号，宾词栏编以“（1）、（2）、（3）、（4）…”为序号。如果各栏间存在一定的计算关系，也可用数学符号表示。例如，（3）=（2）×（1），表示该表中第（3）栏的数字为第（2）栏数字与第（1）栏数字之积。

4. 统计表表式的设计

统计表一般是开口的，即表的左右两侧不画纵线。上下两端的端线以粗线绘制，中间其他线用细线，这样使人看起来清楚、醒目。纵列标题之间一般用竖线隔开，而横行之间通常不必用横线隔开。表的下部如果有合计横行，在其上面用细线将其与其他数字隔开。

5. 指标数值填写的要求

表中的数字应字迹清楚、填写整齐、数位对齐。当数字为 0 时要写出来；当缺某项数字或因数字太小略而不计时，可用符号“…”表示；当没有数字时，应画“—”；当某数与上、下、左、右数字相同时，仍应填写完整，避免出现“同上”、“同左”等字样。表中的数字栏

不应留有空格。

6. 计量单位栏的设计

统计表中必须注明数字资料的计量单位。当全表只有一个计量单位时，可以把它写在表头的右上方。如果表中需要注明不同计量单位，横行的计量单位可以专设一栏；纵栏的计量单位与纵栏标题写在一起，并加括号。

7. 注解或说明的设计

例如，某些指标有特殊的计算口径、某些统计资料只包括一部分地区、某些数字是由估算来插补的等，都要加以说明。为了保证统计数据的科学性与严肃性，必要时还要注明资料的来源，以便查考。注解或说明一般写在表的下方。

二、统计图

（一）统计图的相关认识

1. 统计图的概念

统计图是将抽象的数字绘制成几何图形或象形图形来显示统计资料，用以说明现象数量关系的图形。

统计图是表现统计资料的重要方法之一，它通俗易懂、简明扼要、形象生动，是进行统计研究的科学方法。借助统计图将研究结果表现出来，可以增加读者对统计资料的理解，加深记忆。

2. 统计图的绘制原则

为了使统计图能发挥其应有的作用，在绘制时必须遵守以下原则。

（1）绘制统计图要有明确的目的

统计图是表现统计资料的重要形式之一，是统计整理和统计分析的有效工具。设计和绘制统计图必须有明确的目的性。

（2）绘制统计图要有科学性

统计图和一般美术图画不同，它是实际事物的具体反映，不允许夸张，必须按现象的性质、计算口径、图示目的、绘制方法等方面的严格要求制作，做到准确无误地表现统计资料。

（3）绘制统计图要有艺术性

统计图的布局、形式、线条、字体、色彩等方面力求精心设计、灵活应用，使画面鲜明醒目、生动活泼，以提高图示的效果，加深读者印象。

（4）统计图要适合不同场合使用和不同的对象观看

要根据不同的统计资料和不同的目的绘制不同的图形，每个图形都要有简明扼要的标题，做到内容与形式协调，以提高读者兴趣，吸引读者的注意力。

3. 统计图的绘制步骤

①根据统计研究的目的和要求搜集和选择统计资料。

②按统计资料的性质、内容、分布场合和阅读对象确定统计图形。

③对统计资料进行审核分析、整理和计算，确定绘图比例、线条长度等。

④先绘草图，将统计资料与草图校对准确后绘制正式图形。

⑤书写图名，加注数字、图例及文字说明，如编制单位、日期、资料来源等。

⑥审核图形。审核其是否符合既定目的，图式是否准确、简明、生动；坐标尺度是否妥当；线条、数字、文字是否适用；图号、图例是否齐全。

⑦仔细描绘正式统计图。

⑧全面检查图形，发现错误或遗漏的，予以纠正。

（二）常见的统计图

在实际应用中，常见的统计图主要有直方统计图、扇形统计图、拆线统计图、统计地图等。

统计图一般采用直角坐标系，横坐标用来表示事物的组别或自变量 x，纵坐标常用来表示事物出现的次数或因变量 y；或采用角度坐标（如圆形图）、地理坐标（如地形图）等。其结构包括图名、图目（图中的标题）、图尺（坐标单位）、各种图线（基线、轮廓线、指导线等）、图注（图例说明、资料来源等）等。

1. 直方统计图

用一个单位长度表示一定的数量，根据数量的多少，画成长短相应成比例的直条，并按一定顺序排列起来，这样的统计图称为直方统计图。条形统计图可以清楚地表明各种数量的多少。

直方统计图的特点：

①能够使人们一眼看出各个数据的大小；

②易于比较数据之间的差别。

直方图是统计图资料分析中最常用的图形，如图 3-1 所示。

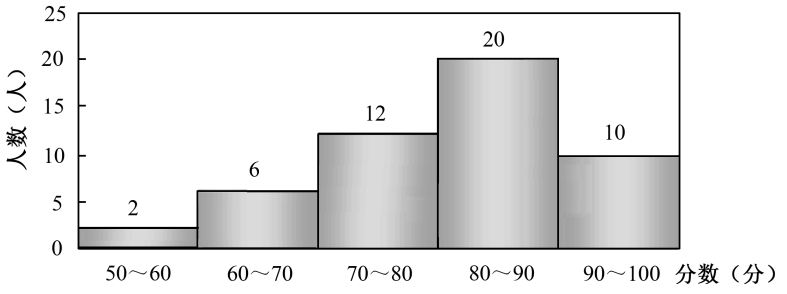


图 3-1　考试成绩分布图

2. 扇形统计图

以一个圆的面积表示事物的总体，以扇形面积表示占总体百分数的统计图叫作扇形统计图，也叫百分数比较图。扇形统计图可以比较清楚地反映出部分与部分、部分与整体之间的数量关系，如图 3-2 所示。

扇形统计图的特点：

①通过扇形的大小来反映各个部分占总体的比例；

②扇形统计图能让读者更清楚地了解各个部分数量同总数之间的关系；

③扇形图可以让一些杂乱无章的数据变得清晰透彻，使人看上去一目了然。

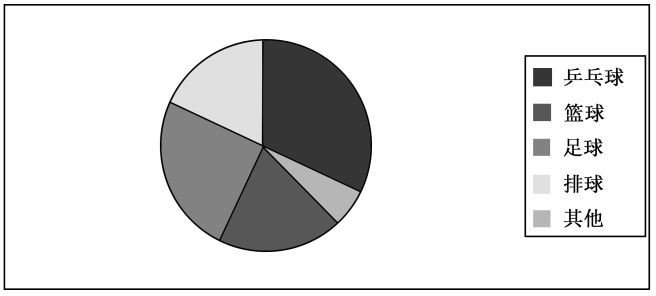


图 3-2　最喜欢的球类活动统计图

3. 折线统计图

以折线的上升或下降来表示统计数量的增减变化的统计图叫作折线统计图。与直方统计图比较，折线统计图不仅可以表示数量的多少，还可以反映同一事物在不同时间里的发展变化情况。折线统计图的特点是能够显示数据的变化趋势、反映事物的变化情况，如图 3-3 所示。

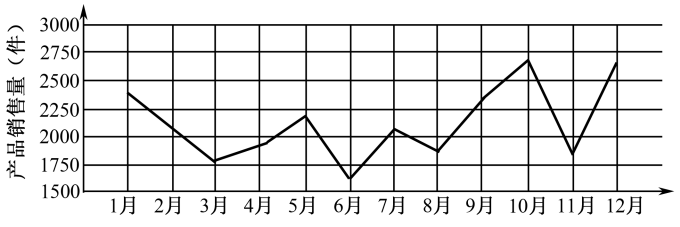


图 3-3　产品销售量折线图

4. 统计地图

统计地图是以地图为背景，利用各种几何图形和不同细纹或颜色，表明和比较事物在各地的分布状况，或对某些现象进行地区间比较的一种统计图形。它一般与统计图形结合使用，用来反映现象的地理位置及与其他自然条件的关系。



本章小结

统计整理是根据统计研究任务的要求，对统计调查所搜集到的原始资料进行科学的分类、汇总，使其条理化、系统化的工作过程。统计分组是统计整理的基础性工作。分配数列是在统计分组的基础上，将总体的所有单位按组归类整理，并按一定顺序排列而形成的总体中各个单位在各组间的分布，又称分布数列或次数分配。

统计表和统计图是通过统计整理得到的反映社会经济现象总体特征的综合资料，是统计工作的初步成果。



复习思考题

1．什么是统计整理？它有何意义？
2．统计整理的步骤有哪些？
3．什么是统计分组？统计分组有哪些作用？
4．什么是分配数列？它有哪些种类？
5．简述统计表的基本结构及分类？
6．选择分组标志应遵循哪些原则？
7．组距式数列所涉及的要素有哪些？



实训基础训练

一、单项选择题

1．采用两个或两个以上标志对社会经济现象总体层叠起来分组的统计方法是（　　）。

A．品质标志分组　　B．复合标志分组
C．混合标志分组　　D．数量标志分组

2．在等距数列中，组距的大小与组数的多少（　　）。

A．成正比　　B．成等比　　C．成反比　　D．不成比例

3．统计分组的核心问题是（　　）。

A．选择分组的标志　　B．划分各组界限

C．区分事务性质　　D．对分组资料再分组

4．说明统计表名称的词句，在统计表中称为（　　）。

A．横行标题　　B．主词　　C．纵栏标题　　D．总标题

5．将 25 个企业按产值分组而编制的变量数列中，变量值是（　　）。

A．产值　　B．企业数　　C．各组的产值数　　D．各组的企业数

二、判断题

1．统计整理是统计工作的最终阶段。（　）

2．设计统计分组的基本原则是，必须保证某一标志在组内各单位之间，以及组与组之间都应具有差异性。（　）

3．统计分组的关键是正确选择分组数列的种类。（　）

4．统计整理阶段最关键的问题是对调查资料进行审核。（　）

5．将总体按照某个标志分组形成的数列都叫变量数列。（　）

6．某一开口式变量数列中其开口组的下限为 200，邻组的组距为 10，则开口组的组中值为 195。（　）

7．在进行分组时，凡遇到某单位的标志值刚好等于相邻两组上、下限数值时，一般是将此值归入上限所在组。（　）

8．变量数列的编制中，在条件不变的情况下，组数分得越多，组距也越大。（　）

三、计算题

1．20 名工人看管机器台数如下：

5　4　2　4　3　4　3　5　4　3

4　3　4　3　2　6　4　4　2　5

试根据上述资料按看管机器台数编制变量数列，并计算出各组频率。

2．某班 45 名学生数学考试成绩如下：

50	65	73	79	90	64	72	78	88
54	66	51	86	76	86	87	77	72
67	74	81	95	91	55	68	74	83
75	75	85	99	84	75	69	58	97
71	62	72	90	78	74	69	82	88

要求：（1）试根据以上资料编制等距数列；

（2）计算组距、组中值；

（3）计算累计次数、累计频率，并绘制分配数列。



课外延伸阅读

我国居民消费价格指数（CPI）是如何调查和生成的

一、关于CPI包含的内容

CPI是英文“Consumer Price Index”的缩写，直译为“消费者价格指数”，在我国通常被称为“居民消费价格指数”。CPI的定义决定了其所包含的统计内容，那就是居民日常消费的全部商品和服务项目，世界各国都采用抽样调查方法进行调查。具体做法就是抽选一组一定时期内居民经常消费的、对居民生活影响相对较大的、有代表性的商品和服务项目，通过调查其价格来计算价格指数，这样既节约了人力，也节省了经费，价格指数也能基本反映居民消费价格的总体变化情况，一举多得。

目前，我国用于计算CPI的商品和服务项目由国家统计局和地方统计部门分级确定。国家统计局根据全国12万户城乡居民家庭消费支出的抽样调查资料统一确定商品和服务项目的类别，设置食品、烟酒及用品、衣着、家庭设备用品及服务、医疗保健及个人用品、交通和通信、娱乐教育文化用品及服务、居住等八大类262个基本分类，涵盖了城乡居民的全部消费内容。

对于升级换代比较快的工业产品，在实际操作中，为保证选取的规格品有代表性，代表规格品一年一定。如果该规格品年中失去代表性或完全从市场上消失，就必须进行更换。在充分听取相关生产企业及销售人员意见的基础上，及时选取另一种有代表性的规格品进行替代，不允许也不会出现长期选用失去代表性或已被淘汰产品的情况。

二、关于价格调查地点的选择

当前，食杂店、百货店、超市、便利店、专业市场、专卖店、购物中心等商业业态众多，农贸市场遍布城乡，这在方便人民群众生活的同时，也使价格调查工作的难度大幅度增加。与不可能调查居民消费的全部商品和服务项目一样，也没必要在每一个销售场所和销售地点都开展价格调查，选取一部分有代表性的商业业态、农贸市场及服务类单位实施抽样调查是最好的选择。

目前，计算我国CPI的价格资料来源于31个省（区、市）共500个调查市县的5万个商业业态、农贸市场，以及医院、电影院等提供服务消费的单位，我们统称为价格调查点。这些调查点主要是依据经济普查获得的企业名录库以及有关部门的行政记录资料，以零售额或经营规模为标志，从高到低排队随机等距抽选出来的，同时按照各种商业业态兼顾、大小兼顾及区域分布合理的原则进行适当调整。

三、关于价格的采集方法

选定代表商品和确定价格调查点以后，接下来的工作就是选择什么方式来收集价格。目前，世界各国根据本国实际通常采用派人直接调查、电话调查、企业报表、网上收集等方式相结合的方法收集计算CPI所需的原始价格资料。1984年，经国务院批准，国家统计局在各

地成立了直属调查队，自此以来一直采用派人直接调查的方式收集原始价格资料，目前分布在 31 个省（区、市）500 个调查市县的价格调查员有 4000 人左右。

为保证源头数据的真实性和可比性，调查员必须要按照统一规范的“三定原则”即“定人、定点、定时”开展价格调查工作。“定人”是指同一个调查员一定时期内固定调查相同的商品项目，目的是让调查员更加专业、更加全面地熟悉和了解这些商品的特征及其属性，避免因不熟悉商品而误将两种不同的商品视为同一种商品的情况发生；“定点”是指固定调查员采集价格的地方（调查点），目的是让调查员熟悉和了解价格调查点的基本情况，便于向销售人员或其他有关人员咨询有效的价格交易信息，准确采集不同采价日同一种商品同一地点的可比价格；“定时”是指固定调查员调查价格的具体时间，保证价格同“时”可比，如调查农贸市场的蔬菜价格，上一次的调查时间是上午 9 时，下一次的调查时间也必须是上午 9 时，若改为 12 时或下午其他时间，这两次的价格就不可比，调查采集的价格就不能用于计算价格指数。

目前，对于 CPI 中的粮食、猪牛羊肉、蔬菜等与居民生活密切相关、价格变动相对频繁的食品，每 5 天调查一次价格；对于服装鞋帽、耐用消费品、交通通信工具等大部分工业产品，每月调查 2~3 次价格；对于水、电等政府定价项目，每月调查核实一次价格。

四、关于权数资料来源

收集齐价格资料以后，就可以计算单个商品或服务项目及 262 个基本分类的价格指数，但还不能计算类别价格指数，还需要各类别相应的权数。在统计中，用来衡量个体对总体作用大小的数值叫作权数。权数一般用百分数（%）或千分数（‰）来表示，又称比重。CPI 中的权数是指每一类别商品或服务项目的消费支出在居民全部商品和服务项目总消费支出中所占的比重。

我国 CPI 中的权数主要是根据全国 12 万户城乡居民家庭各类商品和服务项目的消费支出详细比重确定的。这些资料可以在国家统计局公开编辑出版的有关年鉴中查到。现行制度规定，CPI 中的权数每五年调整一次。但同时也考虑到，随着我国国民经济的持续快速发展，在城乡居民生活水平不断提高的同时，消费结构也在发生变化，加之我们每年都有城乡居民消费支出抽样调查资料，因此每年还要根据全国 12 万户城乡居民家庭消费支出的变动及相关资料对权数进行一次相应的调整。

五、关于 CPI 的汇总计算方法

有了权数资料和基本分类价格指数，就可以计算类别价格指数及 CPI。CPI 的汇总计算方法相对复杂一些，有比较强的专业性。我国采用的汇总计算方法与其他国家基本是一样的。主要区别在于世界上大多数国家仅汇总计算国家一级的 CPI，计算分州、分城市等分区域 CPI 的国家不多。我国的情况相对特殊一些，有分省的 CPI，部分市县也计算 CPI。基本汇总计

算过程如下。

首先，市县统计部门根据国家统计局制定的《流通和消费价格统计调查制度》，按照统一的统计标准、统计口径和计算方法要求，结合当地居民消费的实际情况计算本市县的 CPI。

其次，国家统计局各调查总队对辖区内市县统计部门计算的 CPI 数据进行审核确认后，按人口和消费水平加权汇总计算本省（区、市）的 CPI。

最后，国家统计局对各省（区、市）计算的 CPI 数据进行审核确认后，按人口和消费水平加权汇总计算全国的 CPI。

资料来源：中华人民共和国国家统计局网站统计知识资料汇编



第4章 综合指标

学习目标要求

1. 了解综合指标的各种分类及其表现形式，理解综合指标在认识事物中的作用。
2. 理解总量指标的概念、作用及分类，并能够计算总量指标。
3. 理解相对指标的概念，熟练掌握相对指标的分类及计算方法。
4. 理解平均指标的概念、作用，熟练掌握平均指标的计算方法。
5. 理解标志变异指标的概念、作用，熟练掌握极差、标准差等标志变异指标的计算方法。

重点难点提示

1. 总量指标和相对指标的计算及分析方法，时期指标与时点指标的异同，相对指标的对比关系。
2. 平均指标和标志变异指标的计算及分析方法，权数在计算平均指标中的影响，标准差在标志变异指标中的地位。

教学引导案例

“十二五”时期我国发展取得重大成就

“十二五”时期是我国在实现“两个一百年”奋斗目标历史进程中极为重要的五年。面对错综复杂的国际环境和艰巨繁重的国内改革发展稳定任务，我们党团结带领全国各族人民顽强拼搏、开拓创新，奋力开创了党和国家事业发展新局面，经济社会发展取得了新的重大成就，我国经济实力、科技实力、国防实力、国际影响力又上了一个大台阶。

经济总量稳定增长。“十二五”期间，我国妥善应对国际金融危机持续影响等一系列重大风险挑战，适应经济发展新常态，不断创新宏观调控方式，推动形成经济结构优化、发展动力转换、发展方式转变加快的良好态势。“十二五”规划目标即将胜利实现。2014年与2010年相比，国内生产总值从40.2万亿元增至63.6万亿元；按美元折算，从6万亿美元增至10.4万亿美元，稳居世界第二位；占全球经济比重由9.2%上升到13.3%。2011—2014年我国对世界经济增长的贡献率超过25%。据世界银行数据，我国人均国民总收入由2010年的4300美

元提高至2014年的7380美元，在上中等收入国家中的位次不断上升。对于有13亿人口的大国来讲，这是了不起的成就。

经济结构优化。“十二五”期间，经济结构出现了一些转折性的可喜变化。一是消费对经济增长的拉动作用超过投资，2011—2014年，最终消费对经济增长的年均贡献率为54.8%，高于投资贡献率7.8%。二是第三产业增加值占国内生产总值比重超过第二产业，从44.2%上升到48.1%，比2010年提高3.9%；第二产业占比由46.2%下降到42.7%；第三产业就业占比从34.6%上升到40.6%。三是城镇化率（常住人口）从2010年的49.95%上升到2014年的54.77%。与此同时，农业连续增产增收，粮食产量实现“十一连增”。四是中西部地区投资和经济的增速超过东部地区，地区发展差距扩大的势头得到了初步抑制。

基础设施水平全面跃升。交通、水利、能源、信息等基础设施建设步伐加快。高效、便捷的铁路网、公路网、航空运输网、城际铁路网、航道网逐渐形成。2014年年末，铁路营业里程、公路里程、高速公路里程、定期航班航线里程分别达到11.2万公里、446.4万公里、11.2万公里、463.7万公里，分别比2010年末增长22.6%、11.4%、51%、67.7%。特别是高速铁路运营里程突破1.6万公里，居世界首位。建设了一批跨流域调水和骨干水源工程。信息化水平全面提高，新一代移动通信网、下一代互联网、数字广播电视网、卫星通信等设施建设加速，逐步形成了超高速、大容量、高智能国家干线传输网络；推动了三网互联互通和业务融合。

科技创新能力明显增强。科技整体水平加速提升，一批重大科技成果达到世界先进水平。在基础研究领域，取得量子调控、中微子振荡、90α热休克蛋白、CiPS干细胞等一批重大原创性成果。在载人航天、探月工程、深海潜器、超级计算、北斗导航等战略高技术领域取得重大突破。高铁、4G移动通信、核电、电动汽车、特高压输变电、杂交水稻等重大创新成果加速应用，产生重大效益。2014年，研究与试验发展（R&D）经费支出达13312亿元，占国内生产总值的2.09%，比2010年提高0.36%。科技进步对经济增长贡献率从2010年的50.9%提升到2014年的54%。每万人口发明专利拥有量达到4.9件，比2010年增加2.8件。科技体制改革取得重要突破，企业的技术创新主体地位增强，大众创业、万众创新蓬勃开展。

资料来源：http://gb.cri.cn/42071/2015/11/16/3245s5167936.htm

思考：

1. 主要采用了什么指标描述“十二五”时期我国经济、科技实力显著增强，基础设施水平全面跃升中哪种交通方式里程的增长态势最为明显？

2. 从上述资料来看，能否推导出铁路、公路、高速公路、定期航班航线中，哪种方式里程所占比重较高，哪种方式占主导？

第一节 总 量 指 标

一、总量指标的概念、特点和作用

（一）总量指标的概念

总量指标是反映社会经济现象在一定时间、地点和条件下的总规模、总水平、总成果的

统计指标。它反映被研究对象实在的、绝对的数量，故又称为绝对指标或绝对数。例如，一个国家或地区的人口数、土地面积、粮食产量、国内生产总值、进出口贸易额、商品流转额、固定资产投资额等，都是总量指标。

（二）总量指标的特点

（1）总量指标的表现形式为绝对数，并且要有计量单位。例如，按当年价格计算，2014 我国交通运输、仓储和邮政业增加值为 28 750.0 亿元。

（2）总量指标的数值随着研究范围的大小而增减。例如，一个省的人口一定大于所属的某一个县的人口数；一个县的粮食产量一定大于所属的某一个乡的粮食产量。

（3）只有对有限总体才能计算总量指标。对于无限总体只能采取近似值。

（三）总量指标的作用

1. 总量指标是认识客观现象总体的起点

人们要想全面了解一个客观现象总体的情况，首要问题就是准确地掌握其在一定时间、地点条件下的数量多少和大小。例如，为了科学地指导我国国民经济的协调发展，就必须了解我国人口和劳动力资源、国民财富、土地面积、国民生产总值、国民收入及粮食、钢铁、煤炭、石油的产量等。如果对这些基本的数据都不清楚，那将不能很好地指导经济的发展。

2. 总量指标是进行科学管理的依据

制定政策和编制计划，首先按总量指标形式加以规定。同样，实行社会经济管理，也必须首先从社会经济过程的总量上进行控制。

3. 总量指标是计算相对指标和平均指标的基础

相对指标和平均指标一般是由两个互相联系的总量指标对比计算出来的，它们是总量指标的派生指标。所以总量指标的计算是否科学与合理会直接影响相对指标和平均指标的准确性。

二、总量指标的种类

（一）总量指标按其反映的内容不同，可分为总体单位总量和总体标志总量

1. 总体单位总量

总体单位总量是用来反映总体中总体单位数的多少、说明总体本身规模大小的指标。例如以企业为单位时的企业总数、以职工为单位时的职工总人数、以学校为单位时的学生总人数等。

2. 总体标志总量

总体标志总量是用来反映总体单位某一数量标志所有标志值的总和、表示其数量规模的指标。例如某一总体范围内各企业的增加值、各户粮食产量、某地区用水总量等。

一个总量指标是属于总体单位还是总体标志总量不是固定不变的，随着研究目的的不同和研究对象的变化，它是可以转化的。如果研究对象是企业，则全部企业就是一个总体，它包括的企业个数就是总体单位总量，而企业中的职工人数、产值等指标就是总体标志总量。如果统计的目的在于了解整个职工队伍的收入状况，那么职工总人数就是总体单位总量，而工资总额则是总体标志总量。

（二）总量指标按其反映的时间状况不同，可分为时期指标和时点指标

1. 时期指标

时期指标反映现象在一定时期内发展过程的总量，如产品产量、产值、商品流转额、人口出生数、固定资产投资完成额等。时期指标具有以下特点：①时期指标各时期的数值可以直接相加，其和说明较长时间内社会经济现象发生的总量；②时期指标的数值大小与时期的长短有直接关系，如一年的商品销售额一定大于一个月的商品销售额；③时期指标的数值是通过连续登记取得的。

2. 时点指标

时点指标反映现象在某一时刻（瞬间）状况的总量，如人口数、企业数、生猪存栏头数、固定资产净值、机器台数，商品库存额等。时点指标具有以下特点。①不同时点上的时点指标数值有不具有可加性。不同时点上的时点数值直接相加，除在有关指标的计算过程中需要外，没有实际意义。②时点指标数值大小与时点间隔没有直接关系，如年末的商品库存额不一定比某月月末的库存额大。③时点指标的数值一般是通过间断登记取得的。

（三）总量指标按其采用的计量单位不同，可以分为实物量指标、价值量指标和劳动量指标

1. 实物量指标

实物量指标是以实物单位计量的总量指标，如人口总数、职工人数、粮食产量、钢铁产量、煤炭产量、设备台数等。实物量指标的优点是能直接反映产品的使用价值的大小；其缺点是指标的综合性能较差。

2. 价值量指标

价值量指标以货币单位计量的总量指标，如国民生产总值、国民收入、固定资产投资额、社会商品零售额等。价值量指标的优点是具有广泛的综合性能和概括能力；其缺点是指标比较抽象、脱离物质内容。所以分析问题时总是将实物量指标与价值量指标结合应用。

3. 劳动量指标

劳动量指标是以劳动单位计量的总量指标，如出勤工时、缺勤工时、生产实用工时等。



小贴士

三次产业划分规定

为了更好地反映我国三次产业的发展情况，满足国民经济核算、服务业统计及其他统计调查对三次产业划分的需求，根据《国民经济行业分类》（GB/T 4754—2011）制定本规定。三次产业的范围如下。

第一产业是指农、林、牧、渔业（不含农、林、牧、渔服务业）。

第二产业是指采矿业（不含开采辅助活动），制造业（不含金属制品、机械和设备修理业），电力、热力、燃气及水生产和供应业，建筑业。

第三产业即服务业，是指除第一产业、第二产业以外的其他行业。第三产业包括：批发和零售业，交通运输、仓储和邮政业，住宿和餐饮业，信息传输、软件和信息技术服务业，金融业，房地产业，租赁和商务服务业，科学研究和技术服务业，水利、环境和公共设施管理业，居民服务、修理和其他服务业，教育，卫生和社会工作，文化、体育和娱乐业，公共管理、社会保障和社会组织，国际组织，以及农、林、牧、渔业中的农、林、牧、渔服务业，采矿业中的开采辅助活动，制造业中的金属制品、机械和设备修理业。

（资料来源：http://www.stats.gov.cn/tjsj/tjbz/201301/t20130114_8675.html）

三、总量指标的计算

（一）总量指标的计算方法

总量指标的计算方法主要有两种：直接法和推算法。

1. 直接法

对所有总体单位直接进行调查登记，然后逐步汇总得到总量指标，如月末商品库存量的盘点。

2. 推算法

推算法是根据各种关系推算总量指标或根据非全面调查资料推算总量指标的方法。常用的推算方法有：平衡关系推算法、因素关系推算法、比例关系推算法、插值估算法和抽样推算法等，这些推算方法将在有关章节中讲述。

（二）计算总量指标应注意的问题

1. 必须明确总量指标的计算范围

例如计算交通运输、仓储和邮政业增加值，必须明确交通运输、仓储和邮政业与其他部门的区别，才能得到真实增加值数值。从 1979 年起公路运输包括社会车辆完成数量，从 1984 年起还包括私营运输完成的数量，从 2008 年起公路运输量统计范围原则上为营运车辆。水路运输量统计范围为在交通运输主管部门审批、备案、从事营业性旅客和货物运输生产的船舶。

从 2008 年起，国家铁路和合资铁路客货运输量及周转量统计口径有调整，其中国家铁路包括了国家控股合资部分，合资铁路仅指非控股合资。2010 年、2011 年，合资铁路客运量统计口径和其他年份不一致。

2. 注意现象的同质性

在对实物指标进行汇总时，要注意现象的同质性，不同质的现象水平不能简单相加汇总。

例如，要计算某地区拥有的机动车辆数，要对机动车进行分类（汽车、拖拉机、摩托车等），而后才能统计各类机动车的数量。

3. 计量单位与计算口径要一致

对于一些重要的总量指标，应按照全国统一规定的单位计量。

第二节 相 对 指 标

一、相对指标的概念和作用

（一）相对指标的概念

相对指标又称相对数，是社会经济现象中两个有联系的指标数值之比。其中作为比较基础的量称为基数，作为进行比较的量称为对比数，即

$$\text{相对数}=\frac{\text{对比数}}{\text{基数}}$$

相对指标有以下两个特点：第一，它是一种抽象化了的数值，反映现象之间差别的相对程度；第二，其数值大小不随总体范围的大小而变化。

（二）相对指标的作用

1. 相对指标可以说明事物之间相互联系的程度和事物发展的程度

相对指标有助于人们深入认识事物。例如，我国 2014 年旅客运输量为 2 209 391 万人，这是一个总量指标，是多少呢？仅用这个数值很难对我国的旅客运输量有一个全面认识。那么，怎么办呢？我们可以通过比较来解决这个问题。与 2000 年我国客运量对比 2 209 391/1 478 573＝1.49，说明我国 2014 年客运量与 21 世纪初相比有了显著增长。另外，与其他国家同期的客运量对比，可以反映我国旅客运输与其他国家之间的差异；与计划数对比，可以反映计划完成的情况如何，等等。

2. 相对指标能使一些不宜或不能直接比较的量具备对比的基础，弥补了总量指标的不足

例如，甲厂生产汽车，计划产量 50 万辆，实际产量 60 万辆；乙厂生产冰箱，计划产量 30 万部，实际产量 38 万部。两个企业生产的产品不同，无法用总量指标比较其完成计划的

好坏。通过计算相对数得知，甲企业计划完成程度为 120%，乙企业计划完成程度为 126.67%。可见，乙企业计划完成情况好于甲企业。

（三）相对指标的表现形式

相对指标的表现形式有两种：一是无名数；二是有名数。

1. 无名数

无名数是一种抽象化的数值，具体表现形式有以下两种。

（1）系数或倍数

系数和倍数是将对比基数定为 1 而计算出来的相对数。两个数字对比，分子数值和分母数值差别不大时常用系数，如工资等级系数、固定资产磨损系数和固定资产有用系数等；两个数字对比，分子数值大于分母数值很多时常用倍数，如我国 2014 年客运量为 2 209 391 万人，是 2000 年产量的 1.49 倍。

（2）成数

成数是将对比的基数定为 10 而计算出来的相对数。例如今年粮食产量比去年增加一成，即增产 1/10。

（3）百分数和千分数

百分数和千分数是将对比基数定为 100 或 1000 而计算出来的相对数，分别用%和‰表示。其中百分数是计算相对指标数值时最常用的一种形式，如计划完成程度、发展速度、增长速度都是用百分数表示的。当对比的分子数值比分母数值小很多时，宜用千分数表示，如人口出生率、人口死亡率、人口自然增长率等。

2. 有名数

有名数一般用来表现强度相对指标的数值。它将相对指标中的分子与分母指标数值的计量单位同时使用，以表现事物的密度、强度及普遍程度等。例如人口密度用“人/平方公里”，人均国民生产总值用“元/人”表示等。

二、相对指标的种类

统计中的相对指标根据研究目的和任务的不同，大体上可以分为以下六种。

（一）结构相对指标

结构相对指标是反映同一时期总体内部组成状况的相对数。它是利用分组法，将总体区分为若干部分，以部分数值与总体数值对比计算各部分所占比重的一种相对数，常用百分数的形式表示。其计算公式为：

$$\text{结构相对数}=\frac{\text{总体部分数值}}{\text{总体全部数值}}\times 100\%$$

例如，根据我国六次人口普查资料，总人口性别结构见表 4-1。

表 4-1　人口普查性别结构　　单位：%

	第一次	第二次	第三次	第四次	第五次	第六次
男性	51.82	51.33	51.50	51.60	51.63	51.27
女性	48.18	48.67	48.50	48.40	48.37	48.73
合计	100.00	100.00	100.00	100.00	100.00	100.00

注：数据来源于国家统计局公布的第一次至第六次全国人口普查数据。

由表 4-1 可见，新中国成立以来，我国人口性别结构基本稳定，男性人口在总人口中所占比重略大于女性。

结构相对指标有两个特点：①结构相对指标具有可加性，各部分比重之和等于 1 或百分之百；②结构相对数的分子分母不能互换。

（二）比例相对指标

比例相对指标是反映总体中各部分之间数量联系程度和比例关系的相对数。其计算公式为：

$$\text{比例相对数}=\frac{\text{总体中某一部分的数值}}{\text{总体中另一部分的数值}}$$

在实际中，常见的比例相对指标主要有男女性别比例、农轻重比例、投资和消费比例等。改革开放以来，我国内需总体上保持了较快增长。但是，消费占内需的比重波动下降，投资所占比重则波动上升，投资和消费失衡问题日趋严重。从经验数据出发，考虑到我国目前所处的经济发展阶段，以及各国经验，我国较好的投资与消费比例关系应该是：投资率一般不应高于 40%，消费率不应低于 60%。而且，从历史经验来看，若要保持一个较为合理的投资和消费比例关系，经济增长速度不宜过快。2014 年我国投资率和消费率分别为 45.9%和 51.4%，投资和消费比例为 0.89，与 2000 年的 0.53 相比有明显改善。预计到 2020 年我国投资率可能降到 35%以下，消费率可能上升到 60%以上。

比例相对指标所反映的比例关系属于一种结构性的比例。其作用同结构相对指标相同，只是对比的方法不同，侧重点有所差别。其特点是：①反映的是总体内部各部分之间的比例关系；②分子分母可以互换，其分子和分母既可以是总体单位总量，也可以是总体标志总量；③不具有可加性。

（三）比较相对指标

比较相对指标是反映同一现象在同一时间不同总体的数量对比关系的相对数。其计算公式为：

$$\text{比例相对数}=\frac{\text{某一总体的某类指标数值}}{\text{另一总体的同类指标数值}}$$

这一指标用来说明某类现象在同一时期内不同国家、地区、单位之间的差异程度。例如把企业的各项技术经济指标与同类企业的先进水平对比，或与国家规定的标准对比，可以找出差距，从而为提高企业和整个工业部门的生产水平、经营管理水平提供重要依据。又如，

在表明国家经济实力方面，将反映国家经济力量的主要指标与世界各国的同类指标对比，可以表明国与国之间经济发展水平的差异程度。

2014 年一些国家的人均国内生产总值资料如表 4-2 所示。

表 4-2　2014 年不同国家人均国内生产总值　　单位：美元

国家或地区	人均国内生产总值	国家或地区	人均国内生产总值	国家或地区	人均国内生产总值
中　国	7594	蒙　古	4170	巴　西	11 613
中国香港	40 170	巴基斯坦	1334	委内瑞拉	16 530
中国澳门	96 444	菲 律 宾	2843	捷　克	19 554
孟加拉国	1097	新 加 坡	56 287	法　国	42 736
文　莱	40 776	斯里兰卡	3631	德　国	47 627
柬 埔 寨	1084	泰　国	5561	意 大 利	34 960
印　度	1631	越　南	2052	荷　兰	51 590
印度尼西亚	3515	埃　及	3436	波　兰	14 423
以 色 列	37 032	尼日利亚	3185	俄 罗 斯	12 736
日　本	36 194	南　非	6478	西 班 牙	30 262
哈萨克斯坦	12 276	加 拿 大	50 271	土 耳 其	10 543
韩　国	27 970	墨 西 哥	10 361	乌 克 兰	3082
老　挝	1708	美　国	54 630	英　国	45 603
马来西亚	10 830	阿 根 廷	12 922	澳大利亚	61 887

资料来源：《中国统计年鉴 2015》。

通过表 4-2 可知，2014 年中国人均国内生产总值不足 8000 美元，不仅低于主要发达国家，而且与墨西哥、阿根廷、巴西、委内瑞拉等发展中国家相比差距依然明显。

其特点是：①在一般情况下，分子分母数值可以互换，从不同的角度来说明同一问题；②比较相对数既可以是绝对数对比，也可以是相对数或平均数对比，由于总量指标易受总体范围和生产条件的影响，不能确切地反映要说明的问题，所以计算比较相对数时，更多地采用相对数或平均数进行对比；③比较相对指标不具有可加性。

（四）动态相对指标

动态相对指标是反映同一现象在不同时间上变动程度的相对数。其计算公式如下：

$$\text{动态相对数} = \frac{\text{报告期水平}}{\text{基期水平}} \times 100\%$$

所谓基期，就是用来作为比较标准（基础）的时期；所谓报告期，就是同基期对比的时期，或者说是所研究的时期。

例如，2014 年我国博士研究生招生数为 72 634 人，2013 年为 70 462 人，则：

$$\text{动态相对数} = \frac{72\,634}{70\,462} \times 100\% = 103.1\%$$

计算结果表明，2014 年我国博士研究生招生数与 2013 年相比增加了 3.1%。

动态相对数的特点是：①先发生的作为基期数值，后发生的作为报告期数值；②分子分母不可相互对换。

（五）强度相对指标

强度相对指标是将两个性质不同但有一定联系的总体的指标相对比而得到的相对数。它反映现象的强度、密度和普遍程度。其计算公式如下：

$$\text{强度相对数}=\frac{\text{某一总体的指标数值}}{\text{另一有联系而性质不同的总体的指标数值}}$$

强度相对指标在实际中应用非常广泛，常见的强度相对指标主要有：①反映现象的发展程度、经济实力的强度相对指标，如人均国民生产总值、人均粮食产量、人均能源消耗量等；②反映现象的密度和普遍程度的强度相对指标，如商业网密度、人口密度、铁路或公路网密度等；③反映企业经济效益和经济管理好坏的强度相对指标，如流通费用率、资金利税率、成本利税率、资金产值率、万元 GDP 能耗、万元 GDP 水耗等；④反映社会生产条件或成果的相对指标，如每个职工平均拥有的固定资产额、每万亩耕地拥有的拖拉机台数、每万元产值创造的利润、每千人拥有的医院卫生院床位数、每千人拥有的科技人员数、每千人拥有的大学生人数等；⑤反映社会经济现象变化情况的强度相对指标，如人口出生率、人口死亡率、人口自然增长率等。

强度相对指标的特点是：①强度相对指标在多数情况下是用复名数来表示的，如人口密度用人/平方公里、人均国民生产总值用元/人来表示等，但有时也用千分数或百分数来表示，如人口出生率、死亡率、自然增长率通常用千分数表示，流通费用率用百分数表示；②强度相对指标带有平均的含义，但又区别于平均指标，因为平均指标是将总体标志总量与总体单位总量相比，不涉及两个总体，而强度相对指标则是不同总体间的指标数值之比；③某些强度相对指标的分子分母可以互换，使这一指标有正逆之分，如城市商业网密度，正指标表示每千人拥有的商业网点数，逆指标表示每个商业网点服务的人数。例如，某城市人口为 100 万人，有零售商店 5000 个，则该城市零售商业网密度为：

$$\text{零售商业网点密度}=\frac{5000\text{个}}{1\,000\,000\text{人}}=5\text{个/千人}$$

$$\text{零售商业网点密度}=\frac{1\,000\,000\text{个}}{5000\text{人}}=200\text{个/千人}$$

由此可见，强度相对指标的数值大小与现象的发展程度或密度成正比的叫正指标，与现象的发现程度或密度成反比的叫逆指标。

（六）计划完成情况相对指标

计划完成情况相对指标，是以现象在某一时间内的实际完成数值与计划任务对比而得到的相对数。用来检查、监督计划的完成情况，通常叫计划完成百分比。其基本计算公式如下：

$$\text{计划完成情况相对数}=\frac{\text{实际完成数}}{\text{计划数}}\times 100\%$$

公式中的分母是计划指标数值，分子是反映计划执行结果的实际值。因此，要求分子分母在指标含义、计算口径、计算方法、计量单位、时间长度和空间范围等方面完全一致。

该指标反映了计划的执行情况。其特点是：①由于计划数是衡量计划完成情况的标准，故分子分母不得互换；②判断计划完成程度的好坏，要视指标的类型而定。对于正指标，如产量、产值、劳动生产率等，计划完成情况相对数大于 100%才算超额；对于逆指标，如单位产品成本、流通费用率等，计划完成情况相对数小于 100%才算超额；对于少数指标，如职工人数、工资总额、固定资产投资额等，是不允许突破计划的，这些指标的计划完成情况相对数以 100%为好。

在实际中，计划数不仅可以表现为绝对数，而且可以表现为相对数和平均数，所以，计划完成情况相对指标既可以用绝对数计算，也可以用相对数或平均数计算。

1. 计划数为绝对数时的计算方法

计划数为绝对数时，可以有两种计算方法。

一种是实际完成数和计划数都是同一时期的。它可以用来说明某段时间内计划执行情况的总结果。这种情况可以直接利用基本公式计算。例如，2015 年某企业总产值计划任务数为 4000 万元，实际完成 4200 万元，其计划完成情况为：

$$\text{计划完成情况相对数}=\frac{4200}{4000}\times 100\%=105\%$$

计算结果表明，2015 年该企业超额 5%完成总产值计划。

另一种是自计划期初至某时间的累计完成数对计划期全期计划数之比。它可用来分析整个计划期间计划的执行进度。其计算公式为：

$$\text{计划执行进度相对数}=\frac{\text{自计划期初至某时间的实际累计完成数}}{\text{全期计划数}}\times 100\%$$

例如，某企业 2015 年产值计划完成情况如表 4-3 所示。

表 4-3　某企业 2015 年产值计划完成情况

季度	计划产值（万元）	实际产值（万元）	实际累计产值（万元）	计划执行进度（%）
第一季度	3000	2800	2800	18.1
第二季度	3500	3200	6000	38.7
第三季度	4000	4000	10 000	64.5
第四季度	5000	—	—	—
全　　年	15 500	—	—	—

计算结果表明，该企业 2015 年产值计划执行进度情况是：第一季度只完成全年计划的 18.1%；截止到第二季度完成了全年计划的 38.7%；截止到第三季度，已完成了全年计划进度的 64.5%，因此，该企业应在第四季度采取有针对性的措施，才能完成全年的产值计划。

2. 计划数为相对数的计算方法

计划数为相对数时，计算计划完成情况相对指标的计算公式为：

$$\text{计划完成情况相对数}=\frac{\text{实际完成百分比}}{\text{计划规定手工艺百分比}}\times 100\%$$

它适用于考核社会经济现象的降低率、增长率的计划完成情况。例如考核某物流企业的成本降低率、流通费用降低率、人口自然增长率等计划完成情况。

（1）计划数为增长率时，计划完成情况相对数的计算公式为：

$$\text{计划完成情况相对数}=\frac{1+\text{实际增长率}}{1+\text{计划增长率}}\times 100\%$$

例如，某企业计划规定劳动生产率比上年提高 15%，实际提高了 18%。则：

$$\text{计划完成情况相对数}=\frac{1+18\%}{1+15\%}\times 100\%=\frac{118\%}{115\%}\times 100\%=102.6\%$$

计算结果表明，劳动生产率超额 2.6%完成计划。

（2）计划数为降低率时，计划完成情况相对数的计算公式为：

$$\text{计划完成情况相对数}=\frac{1-\text{实际降低率}}{1-\text{计划降低率}}\times 100\%$$

例如，某物流企业上年度实际人均成本为 500 元，本年度计划比上年降低 4%，实际降低了 5%，则：

$$\text{计划完成情况相对数}=\frac{1-5\%}{1-4\%}\times 100\%=98.96\%$$

计算结果表明，实际成本比计划任务降低了 1.04%。

3. 计划数为平均数时，计算计划完成情况相对指标的公式为：

$$\text{计划完成情况相对数}=\frac{\text{实际平均水平}}{\text{计划平均水平}}\times 100\%$$

它适用于考核以平均水平表示的经济指标的计划完成情况，如劳动生产率、平均成本、平均价格等。

例如，某物流企业单位成本水平计划为 40 元，实际为 45 元，则：

$$\text{计划完成情况相对数}=\frac{45}{40}\times 100\%=113.33\%$$

计算结果表明，产品单位实际成本高于单位计划成本 13.33%。

计划完成情况的检查可分为长期计划执行情况的检查与短期计划执行情况的检查。长期计划执行情况的检查，是指国民经济五年计划或十年计划完成程度的考核，其中，主要是五年计划执行情况的考核。由于我国制定的长期计划有两种情况：一是规定计划期末年应达到的水平，二是规定计划期累计应达到的水平。因而在进行长期计划执行情况检查时有两种方法。

（1）水平法

在下达计划时，如果只规定计划期末应达到的水平，如商品零售额、产品产量等，就可采用水平法检查计划。

所谓水平法，就是用计划期最后一年的实际完成数与计划中规定同期应达到的水平进行对比，以检查计划的完成情况。

$$\text{计划完成情况相对数}=\frac{\text{计划期末年实际达到的水平}}{\text{计划规定的末年水平}}\times 100\%$$

提前完成计划时间＝连续一年的完成数达到计划规定的最后一年的完成数后剩余时间。

例如，某企业计划规定五年计划最后一年的生产量要达到 60 万吨，实际执行情况见表 4-4。

表 4-4　某企业五年期间生产量变动情况　　**单位：万吨**

第一年	第二年	第三年		第四年				第五年			
		上半年	下半年	第一季度	第二季度	第三季度	第四季度	第一季度	第二季度	第三季度	第四季度
47	49	24	26	12	11	14	15	15	16	16	17

试由表 4-4 中的资料计算计划完成程度和提前或推迟完成计划的时间。

$$\text{计划完成程度相对指标}=\frac{\text{实际最末一年水平}}{\text{计划最末一年水平}}\times 100\%$$

$$=\frac{\text{计划期末年实际达到的水平}}{\text{计划规定的末年水平}}\times 100\% =106.7\%$$

提前完成计划时间=12−6=6（个月）

由此可见，该企业超额 6.7%完成计划，提前了六个月即两个季度。

（2）累计法

在下达计划时，如果是规定各年累计应完成的工作量或累计应达到的水平，如新增固定资产、新增生产能力、造林面积等，就应采用累计法检查计划。

所谓累计法，就是用整个计划期间实际完成的累计数和计划数对比，以检查计划的完成情况。

$$\text{计划完成情况相对数}=\frac{\text{计划期间实际累计完成量}}{\text{计划规定手工艺累计量}}\times 100\%$$

提前完成计划时间＝计划期全部时间－自计划执行之日起至累计实际数已达计划数的时间

例如，某五年计划中规定的固定资产投资额为 2 200 亿元，五年内累计实际完成 2 240 亿元，试检查计划完成情况。

又如，假定至最后一年的七月底为止累计投资额已达到 2 200 亿元，问提前完成计划的时间为多少?

$$\text{计划完成情况相对数}=\frac{2\,240\text{亿元}}{2\,200\text{亿元}}\times 100\% =101.82\%$$

提前完成计划时间=60−55=5（月）

三、应用相对指标应注意的问题

相对指标是进行现象数量对比分析的重要工具，正确计算和运用相对指标一般须遵循以

下几个原则。

（一）可比性原则

相对指标是两个相互联系指标的比率，是一个抽象化的数值，因此，用来进行对比的分子和分母必须可比。可比性原则是指对比的两个指标应有相互对比的共同基础，要比得合理，能恰当反映出现象的数量关系，符合研究对象的客观规律。

可比性主要包括两个含义：一是对比的两个指标必须是有联系的；二是对比的两个指标所包括的范围、内容和计算方法等必须一致。只有这样，算出的相对数才有意义，才能正确反映社会经济现象的实质。例如，计算人口密度时，要求人口与土地面积必须都是同一地区的数值，否则是不可比的。又如，在检查计划执行情况时，必须审核实际完成数与计划任务数所包含的指标内容是否一致。

在将我国的统计资料与世界其他国家的统计资料进行对比时，更应该注意指标的内容、范围和计算方法的可比性。例如，我国万元 GDP 能耗的计量单位为标准煤，国际范围内通常用标准油，所以在进行比较时要注意将其折合为可比的指标。

（二）相对指标要与绝对指标结合运用

相对指标往往是由两个有联系的绝对数对比形成的，它把现象的绝对水平抽象化，使许多不便于直接对比的现象可以进行比较，反映出现象之间的相互联系和对比关系。

但是，相对指标是一个抽象化的数值，它只能反映所对比的事物的差别程度，而不能反映这种差别的实际意义，只有将相对指标与总量指标结合运用，才能对所研究的现象有一个比较完整而具体的认识。



小贴士

相对指标与绝对指标

改革开放三十多年来中国经济快速增长，但是经济总量与美国和日本相比仍有较大差距。1979—2007 年，中国国内生产总值年均实际增长 9.8%，不仅明显高于 1953—1978 年平均增长 6.1%的速度，而且大大高于同期世界经济年平均增长 3.0%的速度。

中国 GDP 总量在 2001 年开始超过意大利，升至世界第 6 位，2005 年超过法国，上升至世界第 5 位，2006 年超过英国，升至世界第 4 位，2007 年超过德国，跃居世界第 3 位。但是，从经济总量上来看，根据国际货币基金组织的统计，折合成美元，我国 2007 年国内生产总值为 32 801 亿美元，相当于美国的 23.7%、日本的 74.9%、德国的 99.5%。

与此同时，经济总量占世界经济的份额也有明显上升，1978 年为 1.8%，2007 年提高到 6.0%。在 GDP 净增量方面，2001 年中国仅有 1263 亿美元，2002 年也不过 1290 亿美元，在 2003 年，七国（中国、美国、日本、德国、英国、法国、意大利）GDP 净增量都是正数，中国的净增量排在七国最后。2004 年中国净增量上升到七国第 4 位，仅次于美国、日本和英国。2005 年开始，中国 GDP 净增量仅次于美国。2007 年，中国 GDP 净增量开始超过美国，当年美国 GDP 增量为 6787 亿美元，中国为 7254 亿美元。2008 年，美国 GDP 增量仅为 3638 亿美元，中国为 9454 亿美元。

由此可见，在统计分析中，不能只根据相对数的大小来判断事物的好坏，因为较大的相对数背后的绝对值可能很小，而较小的相对数背后可能隐藏着较大的绝对数。这说明，高速度并不一定代表高水平。相反，低速度代表的水平不一定低。只有把相对指标和总量指标结合起来才能使我们的认识更全面、更正确。

（三）各种相对指标要结合运用

不同的相对指标有不同的作用，运用相对指标进行统计分析时，可以将从不同角度说明问题的各种相对指标结合起来使用，全面地说明客观现象的状况及发展的规律性。

例如，要反映一个国家人口总体的再生产情况，仅计算人口出生率是不够的，必须把人口出生率、死亡率、自然增长率等相对指标结合起来运用，只有这样才能比较全面、准确地反映人口增减变化情况。要反映物流企业的经济效益状况，既要选择反映物流企业财务运行状况的指标，又要选择反映物流企业盈利状况的指标，只有这样才能比较全面、准确地反映物流企业的经济效益情况。

第三节　平 均 指 标

统计数据经过整理后，对数据的分布情况和规律有了直观而形象的大致了解。但是，要想深入地进行分析和研究，还需要找出具有代表性的数量特征值，来准确地描述统计数据的分布特征。统计研究中通常用平均指标反映数据分布的集中趋势。

一、平均指标概述

（一）平均指标的概念

平均指标是反映总体各单位某一数量标志值在具体时间、地点条件下达到的一般水平的综合指标。

例如，我们把成年男人的身高分为 155 厘米以下、155～165 厘米、165～175 厘米、175～185 厘米、185 厘米以上五挡，随机选取一组成年男人，对其身高分布进行观察，落入这五挡的人数往往不相等。但从总体上看，它存在着一种集中趋势，即身高极高和极低的人数较少，多数人的身高为中等。

学生的成绩分布也是如此：分数特别高和特别低的学生总是占少数，多数人的分数总是接近中间的数值。我们常用平均身高来描述人群身高分布的一般水平、用平均分来描述一组学生学习成绩的一般水平。这里的平均身高、平均分是平均指标，它们分别描述了人群身高分布和学生成绩分布的集中趋势。

（二）平均指标的种类

按不同的计算方法，常用的平均指标有：算术平均数、调和平均数、几何平均数、中位数、众数。前三种平均数又称为数值平均数，后两种平均数又称为位置平均数。它们有不同的计算方法，也有不同的应用场合。

（三）平均指标的特点

1. 平均指标必须在同质总体内计算

例如某厂职工的平均工资，只能由该厂范围内的职工人数和相应的工资总额进行计算，凡不属于该厂范围内的职工和相应的工资总额都不能包括在内。

2. 平均指标是一个代表值，它代表的是被研究总体的一般水平

例如，某商场职工的月销售额有高有低，有的职工月销售额为 10 000 元，有的职工月销售额为 16 000 元，有的为 20 000 元，还有的为 28 000 元，等等。若根据该商场各个职工销售额综合计算出职工平均月销售额为 18 000 元，那么 18 000 元就是一个代表值，它反映了该商场职工月销售额的一般水平。

3. 平均指标是一个抽象化了的数值

平均指标把被研究总体各单位的标志值的数量差异抽象化了，这个抽象化的数值实际上并不一定存在，它只说明了总体各单位的某一标志值以平均数为中心上下波动。例如，某商场职工的月平均销售额为 18 000 元，但是各个职工的月销售额有高有低，高于 18 000 元的销售额和低于 18 000 元的销售额互相抵消了，从而得出平均月销售额 18 000 元。所以说，平均指标表明的是数据分布集中趋势的一般特征。

4. 计算平均指标应以大量观察法为基础

只有进行大量观察，将现象偶然性的差异互相抵消，平均指标才能反映出总体内各单位某一数量指标的集中趋势。

（四）平均指标的作用

1. 平均指标可以用来反映数据分布的一般水平

例如职工平均工资反映职工工资收入的一般水平、平均亩产量反映粮食产量的一般水平。

2. 平均指标可以用来对比不同总体的一般水平

平均指标经常用来对同类现象在不同单位、不同地区之间进行比较研究，从而反映现象在不同单位、不同地区之间的差异。

例如，要比较两个企业工人工资水平的高低，仅对比这两个企业的工资总额是不能说明问题的，因为工资总额受到企业职工总人数的影响。更不能用个别工人的工资水平进行比较，因为个别工人工资的多少受多种因素的影响，如工龄、技术水平等。个别工人的工资不能代表企业所有工人的工资水平。只有比较两企业工人的平均工资，才能做出正确的判断。

3. 利用平均指标可以说明事物的发展动态和变化趋势

利用平均指标可以比较和反映现象在同一单位或地区不同时期一般水平的发展变化，说明事物的发展动态和变化趋势。例如，研究某地区粮食生产水平的变化，由于粮食总产量会受播种面积变动的影响，因此以粮食单位面积产量这个平均数进行对比，就能正确地反映出该地区粮食生产水平的发展动态及变化趋势。

4. 平均指标可以作为科学预测与决策的推算依据

平均指标可以作为某些科学预测、决策、管理和某些推算的依据。

二、平均指标的计算与分析

（一）算术平均数

算术平均数是计算平均指标最常用的方法和最基本的形式。这是由于社会经济生活中大量存在的情况是：社会经济现象总体的标志总量为总体各个单位标志值的算术和。例如，企业职工的工资总额是每个职工工资加总而得到的；某校学生总人数是全校各班人数的总和。在这种情况下，平均指标最适合采用算术平均数的形式，基本公式如下：

$$\text{算术平均数}=\frac{\text{总体标志总量}}{\text{总体单位总量}}$$

例如，某企业某月的工资总额为 140 000 元，职工总人数为 100 人，则该企业职工月平均工资为 1400 元。即

$$\text{平均工资}=\frac{\text{工资总额}}{\text{职工总人数}}=\frac{140\,000}{100}=1400(\text{元})$$

由上例可见，计算算术平均数时，总体标志总量依附于总体单位数。使用此公式时，必须保证分子与分母所包含口径严格一致，即各标志值与各单位之间是一一对应的，否则计算平均指标就失去了意义。这也是平均指标与强度相对指标的不同点。

在实际工作中，由于掌握资料不同，算术平均数可分为简单算术平均数和加权算术平均数两种。

1. 简单算术平均数

简单算术平均数就是将总体各单位的标志值简单加总，除以总体单位数而求得的平均数，它适用于未分组资料。

例如，某配送中心某分拣小组有 5 名工人，各人每分钟分拣量为：14、15、16、17、18 件，则这 5 名工人的平均每分钟分拣量为：

$$\text{平均每分钟分拣量}=\frac{14+15+16+17+18}{5}=16(\text{件})$$

将上式用符号表示为：

$$\bar{x}=\frac{x_1+x_2+x_3+\cdots+x_n}{n}=\frac{\sum x}{n}$$

式中，$\bar{x}$ 代表算术平均数；x 代表单位标志值（变量值）；n 代表总体单位数（项数）；$\sum$是总和的符合。



【例 4-1】

某生产小组 10 名工人的日产零件数如下：58，59，60，66，72，75，86，98，109，118。计算该小组工人的平均日产量。

解：该小组工人的平均日产量为：

$$\bar{x}=\frac{58+59+60+66+72+75+86+98+109+118}{10}=80(\text{件})$$

算术平均数的大小只受数据值大小的影响，不受总体中数据个数的影响。

2. 加权算术平均数

当掌握的数据资料已经分组时，可采用加权算术平均数的形式计算平均数。其计算公式为：

$$\bar{x}=\frac{x_1f_1+x_2f_2+\cdots+x_nf_n}{f_1+f_2+\cdots+f_n}$$

式中，$\bar{x}$ 表示加权算术平均数；x_i 表示各组标志值；f_i 表示各组频数，即各组标志值出现的次数。

① 根据单项分组的数据资料计算加权算术平均数。



【例 4-2】

某企业电子产品加工车间的 50 名生产工人日加工零件数情况如表 4-5 所示。计算该车间工人平均日加工零件数。

表 4-5　50 名生产工人日加工零件加权平均数计算表

日加工零件数（件）x_i	工人人数（人）f_i	各组加工零件总数（件）x_if_i
20	5	100
21	10	210
22	10	220
23	20	460
24	5	120
合计	50	1110

解：该车间工人平均日加工零件数为：

$$\bar{x}=\frac{x_1f_1+x_2f_2+\cdots+x_nf_n}{f_1+f_2+\cdots+f_n}=\frac{1110}{50}=22.2(\text{件})$$

按加权算术平均数的方法计算平均数，平均数的大小不仅取决于总体中各单位标志值（x_i）的大小，还受到各标志值出现次数（f_i）多少的影响。标志值出现次数多的组，其标志值对平均数的影响大，标志值出现次数少的组，其标志值对平均数的影响小。标志值次数（f_i）的多少对平均数大小的影响具有权衡轻重的作用。因此，在统计中通常把各组单位数称为权数，把每个标志值乘以权数的过程叫作加权过程。

② 根据组距分组数据资料计算加权算术平均数。

对于组距分组数据资料，要计算加权算术平均数。首先计算各组的组中值，用各组的组中值作为各组的代表值计算平均数。用组中值作为代表值是以每一组内各单位的标志值分布均匀为前提的。因此，用组中值代替各组标志值计算的算术平均数具有一定的假定性，是一个近似值。



【例 4-3】

某电脑公司销售量的数据资料如表 4-6 所示，计算该公司日平均销售量。

表 4-6　某电脑公司日平均销售量计算表

按日销售量分组（台）	组中值（台）x_i	天数（天）f_i	各组销售总量（台）x_if_i
150～160	155	10	1550
160～170	165	25	4125
170～180	175	35	6125
180～190	185	25	4625
190～200	195	5	975
合计	—	100	17 400

解：该电脑公司日平均销售量为：

$$\bar{x}=\frac{x_1f_1+x_2f_2+\cdots+x_nf_n}{f_1+f_2+\cdots+f_n}=\frac{17\,400}{100}=174\ （台）$$

③ 频率为权数的加权算术平均数。

权数对于算术平均数的影响作用，就其实质而言，并不单纯取决于各组单位数次数的多少，而是取决于各组单位数占总体单位数的比重（又称权重系数）的大小，即哪一组单位数所占比重大哪一组标志值对平均数的影响就大。

权数有两种形式：一种是以绝对数表示的，称为次数或频数；另一种是以比重表示的，称为频率。加权算术平均数可采用频率作为权数，同一组数据资料用这两种权数所计算的加权算术平均数完全相同。权数采用频率的形式计算时，加权算术平均数的计算公式为：

$$\bar{x}=x_1\cdot\frac{f_1}{\sum_{i=1}^{n}f_i}+x_2\cdot\frac{f_2}{\sum_{i=1}^{n}f_i}+\cdots+x_n\cdot\frac{f_n}{\sum_{i=1}^{n}f_i}$$

仍以例 4-3 中的资料为例，用频率作为权数计算日平均销售量，计算过程如表 4-7 所示。

表 4-7　某电脑公司日平均销售量计算表

按日销售量分组（台）	组中值（台）x_i	天数（天）f_i	权重系数 $f_i/\sum f_i$	日销售量与频率的乘积 $x_i\cdot(f_i/\sum f_i)$
150 ~ 160	155	10	0.10	15.50
160 ~ 170	165	25	0.25	41.25
170 ~ 180	175	35	0.35	61.25
180 ~ 190	185	25	0.25	46.25
190 ~ 200	195	5	0.05	9.75
合计	—	100	1.00	174

日平均销售量为：$\bar{x}=x_1\cdot\frac{f_1}{\sum_{i=1}^{n}f_i}+x_2\cdot\frac{f_2}{\sum_{i=1}^{n}f_i}+\cdots+x_n\cdot\frac{f_n}{\sum_{i=1}^{n}f_i}=174$（台）。

从以上计算可以看出，用次数权数和频率权数计算的加权算术平均数结果是一致的。由此可见，总体单位数的频率权数是计算加权算术平均数时起作用的实质。当各组频数等比例增加或减少时，仅仅是各组频数发生了变化，而各组的比重（频率）未变，计算的加权算术平均数不会改变。严格地讲，计算加权算术平均数时，其权数应是各组频率，而不是各组的绝对频数。

（二）调和平均数

1. 调和平均数的含义

调和平均数是根据数据分布中各数据值的倒数来计算的算术平均数，所以又称为倒数平均数。在实际应用中，它主要作为算术平均数的变形来使用，它仍然依据算术平均数的基本公式（总体各单位标志值总和除以总体单位总个数）来计算。

2. 调和平均数的计算与应用

根据所掌握的统计资料不同，调和平均数又分为简单调和平均数和加权调和平均数。

（1）简单调和平均数

简单调和平均数适用于对未分组的数据资料计算调和平均数。

简单调和平均数的计算公式为：

$$\overline{x_H}=\frac{1}{\frac{1}{n}\left(\frac{1}{x_1}+\frac{1}{x_2}+\cdots+\frac{1}{x_n}\right)}$$

式中，$\overline{x_H}$ 表示调和平均数；x_i 表示总体各单位标志值；n 表示数据总个数。



【例 4-4】

某市场黄瓜的价格在早、午、晚分别为 2.40 元每千克、1.60 元每千克、1.20 元每千克。如果某顾客早、午、晚各买一元钱的黄瓜，那么黄瓜平均价格是多少？

解：黄瓜的平均价格=$\frac{\text{总金额}}{\text{总购买量}}$。

$$\overline{x_H} = \frac{1}{\frac{1}{n}(\frac{1}{x_1}+\frac{1}{x_2}+\cdots+\frac{1}{x_n})} = \frac{3}{\frac{1}{2.40}+\frac{1}{1.60}+\frac{1}{1.20}} = 1.6\text{（元/千克）}$$

上式分母中的 $\frac{1}{2.40}$、$\frac{1}{1.60}$、$\frac{1}{1.20}$ 分别表示早、午、晚 1 元钱所买到的黄瓜的重量，它们的和表示总购买量；分子 3 表示早、午、晚买黄瓜所用的总金额。这就是简单调和平均数。

若早、午、晚各买 1 千克黄瓜，那么黄瓜的平均价格为：

$$\bar{x} = \frac{x_1+x_2+\cdots+x_n}{n} = \frac{2.40+1.60+1.20}{3} = 1.73\text{（元/千克）}$$

这时采用的是简单算术平均数的方法。

在各买 1 千克的情况下，由于重量相同，每种价格对平均价格的影响是相等的；在各买 1 元钱的情况下，虽然金额相同，但由于价格不同，所以买到的黄瓜的重量就不一样，这时每种价格对平均价格的影响是不相同的。调和平均数就是在这种意义下应用的。

（2）加权调和平均数

加权调和平均数适用于已分组的数据资料。如果已知各组的标志值总数，而不知道各组的数据总个数，则应采用加权调和平均数的方法计算平均数。

加权调和平均数的计算公式为：

$$\overline{x_H} = \frac{m_1+m_2+\cdots+m_n}{\frac{m_1}{x_1}+\frac{m_2}{x_2}+\cdots+\frac{m_n}{x_n}}$$

式中，$\overline{x_H}$ 表示调和平均数；m_i 表示各组标志值总和（权数）；x_i 表示各组标志值。



【例 4-5】

某市场黄瓜的价格在早、午、晚分别为 2.40 元每千克、1.60 元每千克、1.20 元每千克。如果该市场早、午、晚分别卖出 120 元黄瓜、160 元、180 元，那么这一天中黄瓜的平均价格是多少？

解：一天中黄瓜的平均价格为：

$$\overline{x_H} = \frac{m_1+m_2+\cdots+m_n}{\frac{m_1}{x_1}+\frac{m_2}{x_2}+\cdots+\frac{m_n}{x_n}} = \frac{120+160+180}{\frac{120}{2.40}+\frac{160}{1.60}+\frac{180}{1.20}} = \frac{460}{300} = 1.53\text{（元/千克）}$$



【例 4-6】

某公司购进四批同种商品，每批的单位价格及购买金额如表 4-8 所示，求这四批商品的平均价格。

表 4-8 购进商品平均价格计算表

商品	价格（元/件）x	购买金额（元）m	购买量（件）$\frac{m}{x}$
第一批	146	730	5
第二批	152	2280	15
第三批	160	2880	18
第四批	170	1700	10
合计	—	7590	48

表 4-8 中资料的第一、二、三列是给定的，第四列是计算得到的。

解：平均价格 $=\dfrac{\text{购买金额}}{\text{购买数量}}$。

由于每批购买商品的金额不同，所以应采用加权调和平均数的公式计算平均价格。平均价格为：

$$\overline{x_H}=\frac{m_1+m_2+\cdots+m_n}{\frac{m_1}{x_1}+\frac{m_2}{x_2}+\cdots+\frac{m_n}{x_n}}=\frac{7590}{48}=158.13\text{（元/件）}$$

需要说明的是，在例 4-6 中，若已知的是价格和购买量，而缺少购买金额资料，则应采用加权算术平均数公式计算平均价格，见表 4-9。

表 4-9 购进商品平均价格计算表

商品	价格（元/件）x	购买量（件）f	购买金额（元）$m=xf$
第一批	146	5	730
第二批	152	15	2 280
第三批	160	18	2 880
第四批	170	10	1 700
合计	—	48	7 590

表 4-9 中资料的第一、二、三列是给定的，第四列是计算得到的。平均价格为：

$$\bar{x}=\frac{x_1f_1+x_2f_2+\cdots+x_nf_n}{f_1+f_2+\cdots f_n}=\frac{7590}{48}=158.13\text{（元/件）}$$

计算结果与加权调和平均数的计算结果是一样的。

加权调和平均数实际上是加权算术平均数的变形。在研究同一问题时，加权调和平均数同加权算术平均数的实际意义是相同的，只是根据所掌握的资料不同而采用不同的计算方法。在实际应用中，若掌握的是各组标志值和各组频数的资料，则应采用加权算术平均数公式计

算平均数；若掌握的是各组标志值和各组标志值总量而未掌握各组频数，则应采用加权调和平均数公式计算平均数。

以上例题都是对单项分组的数据资料计算加权调和平均数。如果对组距分组数据资料计算加权调和平均数，首先要计算出各组的组中值来代替各组的标志值，然后按加权调和平均数的公式进行计算。

（三）几何平均数

1. 几何平均数的含义

几何平均数是 n 个标志值的连乘积的 n 次方根，适用于计算平均比率和平均速度。

2. 几何平均数的计算与应用

由于掌握的资料不同，几何平均数也可分为简单几何平均数和加权几何平均数。

（1）简单几何平均数

简单几何平均数适用于已知每个比率或每个速度求平均数的情况。

设有 n 个标志值（比率或速度）：$x_1, x_2, \cdots, x_n$，则简单几何平均数的计算公式为：

$$\overline{x_G} = \sqrt[n]{x_1 \cdot x_2 \cdots x_n}$$

式中，$\overline{x_G}$ 表示几何平均数；x_i 表示各个标志值；n 表示标志值个数。

【例 4-7】

某工厂生产某种机械产品，需要经过车工、铣工、钳工和组装四道工序。某月份生产的一批产品在各道工序的合格率分别为 98%、96%、94%、97%，求该批产品的平均合格率。

解：四道工序的平均合格率为：

$$\overline{x_G} = \sqrt[n]{x_1 x_2 \cdots x_n} = \sqrt[4]{0.98 \times 0.96 \times 0.94 \times 0.97} = 96.24\%$$

该批产品的总合格率并不等于各道工序的合格率的总和，后面工序的合格率是在前道工序的制品全部合格的基础上计算的。该批产品的总合格率应等于各道工序合格率的连乘积，所以不能采用算术平均数和调和平均数计算平均合格率，而应采用几何平均数公式求得。

（2）加权几何平均数

加权几何平均数适用于比率或速度已经分组的情况。

设经分组后有 n 个标志值（比率或速度）：$x_1, x_2, \cdots, x_n$，如果各标志值的次数（权数）分别为 $f_1, f_2, \cdots, f_n$，那么加权几何平均数的计算公式为：

$$\overline{x_G} = \sqrt[f_1+f_2+\cdots+f_n]{x_1^{f_1} x_2^{f_2} \cdots x_n^{f_n}}$$

【例 4-8】

某商业银行一笔为期 8 年的建设项目投资，按复利计算收益，投资 8 年的年利率分别是：有 2 年为 5%，有 3 年为 6%，有 3 年为 7%。要求计算这笔投资的年平均利率。

解：按复利规定，前一年的本利是后一年的本金，所以总的本利率等于各年本利率的连

乘积。又由于各种本利率的次数（即年数）不同，所以应采用加权几何平均数公式计算年平均利率。

年平均本利率为：

$$\begin{aligned}\overline{x_G} &= \sqrt[f_1+f_2+\cdots+f_n]{x_1^{f_1}x_2^{f_2}\cdots x_n^{f_n}}\\ &= \sqrt[2+3+3]{1.05^2\times1.06^3\times1.07^3}\\ &= \sqrt[8]{1.6086}\\ &= 1.0612\\ &= 106.12\%\end{aligned}$$

年平均利率为：

$$\overline{x_G}-1=106.12\%-1=6.12\%$$

（四）中位数

1. 中位数的含义

中位数是指将数据分布中的各数据按大小顺序排列后，位于中间位置的那个数据值，用 M_e 表示。

从中位数的定义可知，所研究的数据分布中有一半小于中位数、一半大于中位数。用这样一个中等水平的数据值来反映数据分布的一般水平具有非常直观的代表性。在一些不易计算数值平均数的情况下，可用中位数代表数据分布的一般水平，尤其是在数据分布分散程度比较大时，中位数对数据分布集中趋势的代表性就更强。例如，在职工收入水平差异比较大的单位，要了解职工收入的一般水平，用职工收入分布的中位数作为收入水平的代表值要比用算术平均数更恰当，因为它排除了极端数据的影响。

2. 中位数的计算与应用

（1）对未分组的数据资料确定中位数

对未分组的数据资料确定中位数的方法是：先将数据分布中各数据按大小顺序排列：$x_1\leqslant x_2\leqslant\cdots\leqslant x_n$，则中位数就可以按下面的方式确定：

$$M_e=\begin{cases}x_{\frac{n+1}{2}} & \text{当}n\text{为奇数时}\\ \dfrac{x_{\frac{n}{2}}+x_{\frac{n}{2}+1}}{2} & \text{当}n\text{为偶数时}\end{cases}$$

【例 4-9】

某企业 7 名工人的月劳动生产率分别为 2.2 万元、2.8 万元、2.7 万元、2.6 万元、3.8 万元、3.1 万元、2.9 万元，求这 7 名工人月劳动生产率的中位数。

解：先将这 7 名工人月劳动生产率按大小顺序排列：2.2 万元、2.6 万元、2.7 万元、2.8 万元、2.9 万元、3.1 万元、3.8 万元。

中位数位次为：

$$\frac{n+1}{2}=\frac{7+1}{2}=4$$

即序列中第四个位置上的工人月劳动生产率 2.8 万元为中位数：$M_e=2.8$ 万元。

上例中数据的个数为奇数。如果工人数为 8 人，则数据的个数为偶数。若这 8 人的月劳动生产率分别是：2.2 万元、2.5 万元、2.6 万元、2.7 万元 2.8 万元、2.9 万元、3.1 万元、3.8 万元，则中位数位次为 $\frac{n+1}{2}=\frac{8+1}{2}=4.5$。

该位置在第 4 个位次和第 5 个位次中间，则中位数取第 4 个位次和第 5 个位次对应的数据值的算术平均数。即：

$$M_e=\frac{2.7+2.8}{2}=2.75\text{（万元）}$$

（2）对已分组的数据资料确定中位数

对已分组的数据资料确定中位数，应首先确定中位数所处的位置，然后找出中位数所在的组，再计算中位数的近似值。

①由单项分组数据资料确定中位数。

由单项分组数据资料确定中位数时，应首先计算各组的累计次数（或累计频率），然后按 $\frac{\sum f+1}{2}$ 确定中位数位次，最后根据中位数的位次，将按向上或向下累计的方法累计的累计次数刚好超过中位数位次的组确定为中位数组，该组的标志值即为中位数。

【例 4-10】

某大型商场衬衫部两组营业员月销售衬衫件数资料如表 4-10 所示，求每组营业员月销售衬衫件数的中位数。

表 4-10　商场衬衫部营业员按月销售衬衫件数分组表

按销售衬衫件数分组（件/月）	营业员人数（人）		人数累计（人）	
	甲组	乙组	甲组	乙组
180	3	4	3	4
186	5	5	8	9
187	8	6	16	15
189	7	8	23	23
193	4	3	27	26
195	1	4	28	30
合计	28	30	—	—

解：甲组营业员销售衬衫件数的中位数位次为 $\frac{\sum f+1}{2}=\frac{28+1}{2}=14.5$，说明中位数位于第 14 个与第 15 个营业员之间。根据累计次数可知，第 14 个与第 15 个营业员都在第三组，其标志值相同，都是 187 件，所以，甲组营业员月销售衬衫件数的中位数为 187 件。

乙组营业员销售衬衫件数的中位数位次为 $\frac{\sum f+1}{2}=\frac{30+1}{2}=15.5$，说明中位数位于第 15 个与第 16 个营业员之间。根据累计次数可知，这两个营业员分别属于第三组与第四组，所以应取这两组标志值的算术平均数作为中位数。所以，乙组营业员月销售衬衫件数的中位数为：$\frac{187+189}{2}=188$（件）。

②由组距分组数据资料确定中位数。

由组距分组数据资料确定中位数的方法是：首先根据累计次数 $\frac{\sum f+1}{2}$ 确定中位数所在的组，这个组的上下限就规定了中位数可能的取值范围。然后，在假定中位数所在组内各单位均匀分布的前提下，按下列公式计算中位数的近似值。

根据组距分组数据资料计算中位数，有下限公式和上限公式两种。

下限公式：$M_e=L_{M_e}+\frac{\frac{\sum f}{2}-S_{M_e-1}}{f_{M_e}}\times i$。

上限公式：$M_e=U_{M_e}-\frac{\frac{\sum f}{2}-S_{M_e+1}}{f_{M_e}}\times i$。

式中，M_e 为中位数；L_{M_e} 为中位数所在组的下限；U_{M_e} 为中位数所在组的上限；f_{M_e} 为中位数所在组的次数；S_{M_e-1} 为向上累计至中位数所在组前一组止的累计次数；S_{M_e+1} 为向下累计至中位数所在组前一组止的累次次数；i 为中位数所在组的组距。

下限公式适用于以向上累计的方法计算累计次数的情况。上限公式适用于以向下累计的方法计算累计次数的情况。



【例 4-11】

某城市居民家庭年收入情况的抽样调查资料如表 4-11 所示，求居民家庭年收入的中位数。

表 4-11 某城市居民家庭年收入情况

家庭年收入（元）	居民户数（户）	累计次数（户）	
		向上累计	向下累计
15 000 以下	50	50	1000
15 000～20 000	120	170	950
20 000～25 000	230	400	830
25 000～30 000	310	710	600
30 000～35 000	170	880	290
35 000～40 000	80	960	120
40 000 以上	40	1000	40
合计	1000	—	—

解：先计算中位数的位次为：$\frac{\sum f+1}{2}=\frac{1000+1}{2}=500.5$。

再计算出累计次数，见表 4-11 中的第三列和第四列，观察发现，无论是向上累计还是向下累计，第 500.5 项都包括在第 4 组中，即位于 25 000～30 000 这一组，因此 25 000～30 000 这一组就是中位数所在组。再根据公式确定中位数的近似值。

按下限公式计算中位数：

$$M_e=L_{M_e}+\frac{\frac{\sum f}{2}-S_{M_e-1}}{f_{M_e}}\times i=25\,000+\frac{\frac{1000}{2}-400}{310}\times 5000=26\,612.90\ （元）$$

按上限公式计算中位数：

$$M_e=U_{M_e}-\frac{\frac{\sum f}{2}-S_{M_e+1}}{f_{M_e}}\times i=30\,000-\frac{\frac{1000}{2}-290}{310}\times 5000=26\,612.90\ （元）$$

需要明确的是，当累计次数是向上累计时，则要用下限公式计算中位数；当累计次数是向下累计时，则要用上限公式计算中位数。一般情况下，多用下限公式计算中位数。

（五）众数

1. 众数的含义

众数是指一组数据中出现次数最多的标志值，用 M_o 表示。

由于众数是出现最频繁的数据值，因此，在某些情况下它比算术平均数更能鲜明地反映出数据分布的特征和集中趋势。例如，为了掌握集贸市场上某种商品的价格水平，不必登记全部的成交量和成交额，只需用该市场上这种商品最普遍的成交价格代表。

假定市场上这种商品最多的成交价格为 3.8 元每千克，则 3.8 元每千克就可以用来代表这种商品的一般价格水平。显然，用众数来表现数据分布的集中趋势具有非常直观的代表意义。不仅可以对数量标志的数据资料计算众数，还可以对品质标志的数据资料计算众数。例如，销售量最多的服装款式或色彩，即通常所谓的“流行款式”，就属于这种意义的众数。

2. 众数的计算与应用

根据所掌握的资料不同，确定众数的方法分为以下两种。

（1）对未分组的数据资料确定众数

对未分组的数据资料确定众数的方法：数据分布中出现次数最多的数值就是众数。



【例 4-12】

12 名学生参加某门课程的考试，成绩分数分别为：56，61，75，66，76，82，76，93，88，76，77，89。求这组数据的众数。

解：在这组分数中，出现次数最多的是 76 分，所以众数 M_o=76（分）。

需要特别说明的是：如果各标志值分布很均匀，即每个标志值出现的次数相同，则这组数据没有众数。

【例 4-13】

有 20 名学生参加某门课程的考试，所得的分数均不相同，资料如下：56，61，63，64，65，66，68，70，71，72，74，76，77，78，79，80，83，84，88，90，求这组数据的众数。

解：这 20 个分数出现的次数都是 1 次。由这 20 名学生的考试分数所组成的这个数列就没有众数。

（2）对已分组的数据资料确定众数

①在单项分组的数据分布中，出现次数最多的标志值即为众数。

【例 4-14】

某车间 50 名工人每天加工零件情况如表 4-12 所示，计算众数。

表 4-12 某车间工人日加工零件数统计表

日加工零件数（件）	工人数（人）
12	3
14	5
16	32
17	6
18	4
合　计	50

解：从表 4-12 中可直接看出，每天加工零件 16 件的工人最多，那么 16 件就是众数，M_o=16 件。

②根据组距分组的数据资料确定众数。

根据组距分组的数据资料确定众数，通常利用插补法来确定众数的近似值，需要分两步进行。

第一步，确定众数组，即从数据分布中找出次数或频率最大的组，该组的上限、下限就规定了众数可能的取值范围。

第二步，依据与众数组相邻的两个组的次数，利用公式计算众数的近似值。

众数的计算公式如下。

下限公式：$M_o = L_{M_O} + \dfrac{f_{M_o} - f_{M_O-1}}{(f_{M_O} - f_{M_O-1}) + (f_{M_O} - f_{M_O+1})} \times i$。

上限公式：$M_o = U_{M_O} - \dfrac{f_{M_o} - f_{M_O+1}}{(f_{M_O} - f_{M_O-1}) + (f_{M_O} - f_{M_O+1})} \times i$。

式中，M_o 为众数；L_{M_O} 为众数组下限；U_{M_O} 为众数组上限；f_{M_O} 为众数组次数；f_{M_O-1} 为众数组前一组次数；f_{M_O+1} 为众数组后一组次数；i 为众数组组距。



【例 4-15】

根据表 4-11 所给出的某城市居民家庭年收入情况的抽样调查资料，求居民家庭年收入的众数。

解：表中将全部被调查居民户按收入高低分为 7 组，并列出了各组的户数和累计户数资料。容易看出，表中第 4 组（年收入为 25 000～30 000 元的组）是该变量数列的众数组。

按下限公式计算众数：

$$\begin{aligned} M_o &= L_{M_O} + \frac{f_{M_o} - f_{M_O-1}}{(f_{M_O} - f_{M_O-1}) + (f_{M_O} - f_{M_O+1})} \times i \\ &= 25\,000 + \frac{310 - 230}{(310 - 230) + (310 - 170)} \times (30\,000 - 25\,000) \\ &= 26\,818.18(\text{元}) \end{aligned}$$

按上限公式计算众数：

$$\begin{aligned} M_o &= U_{M_O} - \frac{f_{M_o} - f_{M_O+1}}{(f_{M_O} - f_{M_O-1}) + (f_{M_O} - f_{M_O+1})} \times i \\ &= 30\,000 - \frac{310 - 170}{(310 - 230) + (310 - 170)} \times (30\,000 - 25\,000) \\ &= 26\,818.18(\text{元}) \end{aligned}$$

该城市居民家庭年收入的众数为 26 818.18 元。

第四节　标志变异指标

平均数是一组数据一般水平的代表值、中心值，说明了数据分布的集中趋势。但仅用平均数来描述统计数据的分布特征是远远不够的。例如，有一个班的学生某次测验分数如下。

女同学：73，75，77，78，80，82，82，83，85，85。

男同学：50，63，69，74，80，86，93，94，95，96。

这两组学生的平均分都是 80 分，但两组成绩的分散程度不同。女同学的成绩比较集中、整齐，即变异较小，从而平均数的代表程度较高；男同学的成绩比较分散，参差不齐，即变异较大，平均数的代表程度也就较低。平均数在反映数据分布一般水平的同时，掩盖了分布中各数据的数量差异。标志变异指标恰好弥补了这方面的不足，它综合反映了数据分布中各数据值的差异程度，说明了平均数对所平均的数据资料的代表性，也说明了数据分布的特征。

一、标志变异指标的概念

标志变异指标是用来综合反映一组数据中各数据值偏离其中心值（平均数）的趋势的指标。常用的标志变异指标有极差、平均差、方差、标准差、标准差系数等。

在对数据分布进行综合分析时，将平均指标和标志变异指标互相配合、互相补充，可以对数据分布进行较全面的分析。

二、标志变异指标的作用

1. 标志变异指标用来说明数据分布的离散程度

标志变异指标可以用来说明数据分布的离散程度。标志变异指标值越大，说明数据分布的离散程度越大；标志变异指标值越小，说明数据分布的离散程度就越小。

2. 标志变异指标是衡量平均数代表性的尺度

平均数作为数据分布一般水平的代表值或中心值，其代表性强弱与数据分布的离散程度有直接关系。通常情况下，数据分布的标志变异指标值越大，平均数的代表性越小；反之，标志变异指标值越小，平均数的代表性越大。

例如前面给出的一个班学生某次测验分数的资料中，男生、女生的平均分都是 80 分，但各组的分散程度不同，意味着平均数的代表性不同。男生的成绩在平均数 80 分周围分布较为分散，各分数值与平均分的离差较大，标志变异指标值就大，这时平均分代表性较低；女生的成绩在平均数 80 分周围分布较为集中，各分数值与平均分的离差较小，标志变异指标值也较小，这时平均分代表性较高。

3. 标志变异指标是反映社会经济活动过程均衡性或节奏性的重要指标

在经济发展过程中出现的上下波动、升降起伏较大的非均衡变化的现象，以及生产过程中出现的前紧后松或前松后紧的无节奏状况等，都可以利用标志变异指标进行测定和分析。如果标志变异指标值较小，说明现象的发展比较平稳，其均衡性、稳定性、节奏性较好；反之标志变异指标值较大，则表明现象发展的稳定性较差。

4. 标志变异指标可以用来确定统计推断的准确程度

在抽样调查中，根据样本指标来推断总体指标时，标志变异指标值是确定统计推断准确程度及误差大小不可缺少的重要资料。

标志变异指标是统计分析的基本指标，在相关分析、趋势分析、抽样推断和预测决策中都要用到这一指标。

三、标志变异指标的计算与分析

（一）极差

1. 极差的含义

极差是指数据分布中两个极端值（最大值和最小值）之差，又称全距，用 R 表示。

2. 极差的计算与应用

计算极差是测定数据分布离散程度最简单的方法，极差的大小能说明数据值变动范围的大小。极差的计算公式为：

$$R=\text{最大数据值}-\text{最小数据值}$$

根据所掌握的资料不同，极差的计算分为两种情况。

（1）对未分组或单项分组数据资料计算极差

对未分组或单项分组数据资料计算极差，只需用公式 $R=$最大数据值−最小数据值即可求得。

仍以上述测验成绩为例，女同学得分的极差 $R_1=85-73=12$（分），男同学得分的极差 $R_2=96-50=46$（分）。说明男同学测验成绩的离散程度比较大，平均分 80 分的代表性较差。

（2）根据组距分组数据资料计算极差

根据组距分组数据资料求极差的计算公式为：

$$R=\text{最高组上限}-\text{最低组下限}$$

极差主要用来说明数据分布中各数据值变动的总范围。极差越大，说明数据分布中各数据值变动范围越大；反之，极差越小，说明各数据值变动范围越小。在两个总体平均水平相同的情况下，全距越小，平均数的代表性越大；全距越大，平均数的代表性越小；极差为 0 时，平均数的代表性最好。



小贴士

极差的优点和缺点

极差的优点是计算方法简单。极差经常用于检查产品质量的稳定性或进行质量控制。在仅需要了解数据分布离散的范围时，极差是一个比较适用的指标。

极差的缺点是它的大小仅取决于两个极端值，没有考虑到中间各个数值的分布情况，而且当数据分布有异常值存在时，会直接影响极差的大小，就使得极差不能充分反映数据分布中所有各项数据的离中趋势。极差是一个较粗糙的测定数据分布离中趋势的指标。

（二）标准差

1. 标准差的含义

标准差是数据分布中所有数据值与其平均数的离差平方的平均数的平方根，又称均方差，用 σ 表示。标准差的平方就是方差，用 σ^2 表示。

方差的计量单位不便于从经济意义上进行解释，所以实际统计工作中多用标准差来描述统计数据的离散程度。标准差是测定数据分布离散程度最主要的指标。标准差的实质是各数据值对其算术平均数的平均距离。由于标准差是采用平方的方法来消除各数据值与算术平均数离差的正负号，因此它的数学性质比较理想。

标准差越大，表明数据分布的离散程度越大；标准差越小，表明数据分布的离散程度越小。

在两个总体的平均水平相同的情况下，标准差越小，平均数的代表性越大；标准差越大，平均数的代表性越小。

2. 标准差的计算与应用

根据所掌握的资料不同，标准差的计算分为简单标准差和加权标准差两种形式。

（1）简单标准差

对于未分组的数据资料，计算简单标准差，其计算公式为：

$$\sigma=\sqrt{\frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}{n}}$$

为了使计算过程更加简化，标准差也可以用以下公式计算：

$$\sigma=\sqrt{\frac{\sum_{i=1}^{n}x_i{}^2}{n}-\bar{x}^2}$$



【例 4-16】

根据表 4-13 所给工资资料（未分组资料）计算某公司两个部门员工工资的标准差。结果如表 4-14 所示。

表 4-13　两部门员工工资表

甲部门员工工资（元）	乙部门员工工资（元）
1050	950
1150	1100
1200	1200
1250	1300
1350	1450

解：先计算甲部门员工的平均工资：$\bar{x}_{甲}=1200$ 元，乙部门员工的平均工资：$\bar{x}_{乙}=1200$ 元。

表 4-14　两部门员工工资标准差计算表

甲部门（$\bar{x}_{甲}=1200$ 元）			乙部门（$\bar{x}_{乙}=1200$ 元）		
员工工资（元）x	$x-\bar{x}$	$(x-\bar{x})^2$	员工工资（元）x	$x-\bar{x}$	$(x-\bar{x})^2$
1050	−150	22 500	950	−250	62 500
1150	−50	2500	1100	−100	10 000
1200	0	0	1200	0	0
1250	50	2500	1300	100	10 000
1350	150	22 500	1450	250	62 500
合　计	—	50 000	合　计	—	145 000

甲部门员工工资的标准差为：

$$\sigma_{甲}=\sqrt{\frac{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2}{n}}=\sqrt{\frac{50\,000}{5}}=100\text{（元）}$$

乙部门员工工资的标准差为：

$$\sigma_{乙}=\sqrt{\frac{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2}{n}}=\sqrt{\frac{145\,000}{5}}=170.29\text{（元）}$$

计算结果表明，该公司甲、乙两部门员工的平均工资都是 1200 元，但甲部门员工工资的标准差小于乙部门员工工资的标准差，说明甲部门员工工资分布比较均匀，乙部门员工工资分布比较分散，甲部门员工平均工资的代表性比乙部门大。

（2）加权标准差

对于已分组的数据资料，计算加权标准差，其计算公式为：

$$\sigma=\sqrt{\frac{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2 f_i}{\sum_{i=1}^{n}f_i}}$$

为了使计算过程更加简化，加权标准差与方差也可以用以下公式计算：

$$\sigma=\sqrt{\frac{\sum_{i=1}^{n}x_i^2 f_i}{\sum_{i=1}^{n}f_i}-\overline{x}^2}$$

【例 4-17】

某班级统计学期末考试成绩如表 4-15 第一列和第二列所示，计算该班学生成绩的标准差。

表 4-15　某班级“统计学”期末成绩标准差计算表

成绩（分）	人数（人） f	组中值 x	组总分数 xf	离差 $x-\overline{x}$	离差平方加权 $(x-\overline{x})^2 f$
60 以下	2	55	110	−25	1250
60～70	5	65	325	−15	1125
70～80	16	75	1200	−5	400
80～90	20	85	1700	5	500
90～100	7	95	665	15	1575
合　计	50	—	4000	—	4850

解：

$$\overline{x}=\frac{\sum_{i=1}^{n}x_i f_i}{\sum_{i=2}^{n}f_i}=\frac{4\,000}{50}=80\text{（分）}$$

$$\sigma=\sqrt{\frac{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2 f_i}{\sum_{i=1}^{n}f_i}}=\sqrt{\frac{4\,850}{50}}=9.85\text{（分）}$$

该班学生统计学期末考试成绩的标准差为 9.85 分。

（三）标准差系数

1. 标准差系数的含义

极差和标准差都是以绝对数或平均数表示的测定数据分布标志变异的指标。这些指标都带有与数据分布中各数据相同的计量单位，并且其大小与数据分布的平均水平有关。因此在对两个不同总体的离散程度进行比较时，如果这两个总体的平均数相差较大或计量单位不同，就不能采用这类指标直接比较。

为了对比分析平均水平不同或计量单位不同的两个总体的离散程度，必须消除平均数不同或计量单位不同带来的影响，这就需要用极差、标准差与相应的平均数之比来表明数据分布的离散程度。我们将极差、标准差与相应平均数之比分别称为极差系数、标准差系数，统称为离散系数。标准差系数在实际中应用最为广泛，这里主要介绍标准差系数。

2. 标准差系数的计算与应用

标准差系数是一组数据的标准差与其算术平均数之比，用符号 V_σ 表示。其计算公式为：$V_\sigma=\frac{\sigma}{\overline{x}}\times100\%$。



【例 4-18】

甲、乙两个商场营业员月平均销售额分别为：$\overline{x_{甲}}=20000$ 元，$\overline{x_{乙}}=12000$ 元；标准差分别为 $\sigma_{甲}=600$ 元，$\sigma_{乙}=500$ 元。计算标准差系数，并分析哪个商场营业员的月销售额情况比较均匀、哪个商场营业员月平均销售额更具代表性。

解：甲商场营业员月销售额的标准差系数为：

$$V_{\sigma甲}=\frac{\sigma}{\overline{x}}\times100\%=\frac{600}{20000}\times100\%=3.00\%$$

乙商场营业员月销售额的标准差系数为：

$$V_{\sigma乙}=\frac{\sigma}{\overline{x}}\times100\%=\frac{500}{12\,000}\times100\%=4.17\%$$

虽然甲商场营业员月销售额的标准差大于乙商场，但由于两商场营业员月平均销售额相差较大，不能由此说明甲商场营业员的月平均销售额的代表性比乙商场的小。为了说明两商场营业员月销售额的离散程度，可以通过比较标准差系数实现。由计算结果$V_{\sigma甲}<V_{\sigma乙}$可以得出结论，甲商场营业员的月销售额情况比乙商场均匀，甲商场营业员月平均销售额更具有代表性。

标准差系数是无名数，在应用时不受计量单位的限制，适用于对比分析平均水平不同或计量单位不同的两组数据的离散程度的大小。标准差系数大说明该组数据的离散程度大，标准差系数小说明该组数据的离散程度小。



本章小结

总量指标是反映在一定时间、空间条件下某种现象的总体规模、总水平或总成果的统计指标，也称为绝对数。相对指标是两个有联系的指标数值进行对比计算的结果，又称为相对数。平均指标反映的是总体各单位某一数量标志值在具体时间、地点条件下达到的一般水平，反映了现象的集中趋势。标志变异指标是反映总体单位标志值的变动幅度或离差程度的综合指标，标志变动指标越大，说明总体各单位标志值差异程度越大。测定标志变异的指标主要有极差、标准差和标准差系数。



复习思考题

1．什么是总量指标？它有什么作用？

2．相对指标有哪几种？各种相对指标如何计算？计算和应用相对指标应注意哪些问题？

3．什么是平均指标？各有什么作用？使用平均数时应遵守哪些规则？

4．什么是中位数和众数？它们在统计应用中有哪些用途？

5．什么是标志变异指标？它有哪些作用？

6．时期指标和时点指标各有哪些特点？



实训基础训练

一、单项选择题

1. 在不掌握各组频数资料，只掌握各组标志值和各组标志总量的情况下，宜采用（　　）。

A．加权算术平均数　　B．加权调和平均数

C．几何平均数　　D．简单算术平均数

2．总量指标是用（　　）表示的。

A．绝对数形式　　B．相对数形式　　C．平均数形式　　D．百分数形式

3．在下列平均指标中，不受极端值影响的平均指标是（　　）。

A．中位数　　B．算术平均数　　C．调和平均数　　D．几何平均数

4．下面属于时期指标的是（　　）。

A．商场数量　　B．营业员人数　　C．商品价格　　D．商品销售量

5．标志变异指标中，由数据分布中最大值和最小值之差决定的是（　　）。

A．标准差系数　　B．标准差　　C．平均差　　D．极差

6．某地区有 30 万人口，共有 20 个医院，平均每个医院要服务 15 000 人，这个指标是（　　）。

A．平均指标　　B．强度相对指标　　C．比例相对指标　　D. 比较相对指标

7．已知两个同类型企业职工工资的标准差系数分别为 $V_{\sigma甲}$=4.2%和 $V_{\sigma乙}$=4.8%，则两个企业职工平均工资的代表性是（　　）。

A．甲大于乙　　B．乙大于甲　　C．甲乙相同　　D．无法判断

8．2008 年甲省 GDP 是乙省的 134.18%，该指标是（　　）。

A．强度相对指标　　B．动态相对指标　　C．比较相对指标　　D. 比例相对指标

二、判断题

1．中位数和众数都属于平均指标，因此它们数值的大小受到总体内各单位标志值大小的影响。（　）

2．标志变异指标数值越大，说明总体中各单位标志值的变异程度越大，则平均指标的代表性越小。（　）

3．同一个总体，时期指标值的大小与时期长短成正比，时点指标值的大小与时点间隔成反比。（　）

4．标准差是数据分布中所有数据值与其平均数的离差平方的平均数的平方根，它反映数据分布的离中趋势。（　）

5．某年甲、乙两城市社会商品零售额之比为 1:3，这是一个比例相对指标。（　）

6．当各组次数相等时，加权算术平均数等于简单算术平均数。（　）

7．计划完成相对指标的数值大于 100%，就说明完成并超额完成了计划。　（　）

8．算术平均数的大小只受总体各单位标志值大小的影响。　（　）

三、计算题

1．某企业 2012 年计划产值比 2011 年增长 6%，实际完成计划的 110%，问该企业 2012 年产值计划超额完成多少？又知该企业产品每件成本计划在 2011 年 750 元的水平上降低 25 元，实际 2012 年每台成本是 710 元，试计算降低成本计划完成程度指标。

2．某企业下属三个分公司营业收入完成情况如表 4-16 所示，请用相对指标的原理计算并填写表中空缺数值。

表 4-16　各公司营业收入情况表

	去年完成营业收入（万元）	今　年				今年为去年的（%）
		计　划	实　际		计划完成相对数（%）	
		营业收入（万元）	营业收入（万元）	比重（%）		
一公司			1067		97	106.7
二公司	1200	1320			100	
三公司	1800	1890				115.5
合计						

3．甲公司五年计划规定其生产的某种工业产品在该五年计划的最后一年生产量应达到 700 万吨，实际生产情况如表 4-17 所示。

表 4-17　某工业产品生产量情况表　　单位：万吨

年份	2012	2013	2014	2015				2016			
				第一季度	第二季度	第三季度	第四季度	第一季度	第二季度	第三季度	第四季度
产量	450	480	520	160	160	180	170	180	200	210	210

计算该产品计划完成情况，提前多长时间完成？



课外延伸阅读

中国物流业的发展简况与行业地位

中国的“物流”是从日文引进来的外来词，源于日文资料中对“Logistics”一词的翻译“物流”。中国的物流术语标准将物流定义为：物流是物品从供应地向接收地的实体流动过程中，根据实际需要，将运输、储存、采购、装卸搬运、包装、流通加工、配送、信息处理等

功能有机结合起来实现用户要求的过程。

现代物流是经济全球化的产物，也是推动经济全球化的重要服务业。物流业是融合运输业、仓储业、货代业和信息业等的复合型服务产业，是国民经济的重要组成部分，涉及领域广，吸纳就业人数多，促进生产、拉动消费作用大，在促进产业结构调整、转变经济发展方式和增强国民经济竞争力等方面发挥着重要作用。世界现代物流业呈稳步增长态势，欧洲、美国、日本成为当前全球范围内的重要物流基地。中国物流行业起步较晚，随着国民经济的飞速发展，中国物流行业保持较快增长速度，物流体系不断完善，行业运行日益成熟和规范。

2013 年全国社会物流总额 197.8 万亿元，名义同比增长 11.6%，按可比价格计算，同比增长 9.5%，增幅比上年回落 0.3%。分季度看，1 季度增长 9.4%，上半年增长 9.1%，前三季度增长 9.5%，呈现由“稳中趋缓”向“趋稳回升”转变的态势，全年总体保持较快增长。

从构成情况看，2013 年，工业品物流总额 181.5 万亿元，同比增长 9.7%，增幅比上年回落 0.3%。进口货物物流总额 12.1 万亿元，同比增长 6.4%，增幅比上年回落 1.3%。农产品物流总额同比增长 4.0%，增幅比上年回落 0.6%。受电子商务和网络购物快速增长带动，单位与居民物品物流总额保持快速增长态势，同比增长 30.4%，增幅比上年加快 6.9%；受绿色经济、低碳经济和循环经济快速发展带动，再生资源物流总额快速增长，同比增长 20.3%，增幅比上年加快 10.2%。2013 年，社会物流总费用 10.2 万亿元，同比增长 9.3%，增幅较上年同期回落 2.1%。2013 年，社会物流总费用与 GDP 的比率为 18.0%，与上年基本持平。以上数据反映出中国经济社会运行的物流成本仍然较高。

物流产业对国民生产总值的贡献程度可通过物流产业所创造的产值占国民生产总值的比例来衡量。2001—2008 年，物流业增加值占 GDP 的比重呈波动下降趋势，2009—2013 年呈波动上升趋势，说明物流业在整个国民经济中的地位有所提高。2013 年，中国物流业增加值 3.9 万亿元，同比增长 8.5%，增幅较上半年回升 1.1%，但较上年同期回落 0.7%。物流业增加值占 GDP 的比重为 6.8%，占服务业增加值的比重为 14.8%。

资料来源：http://baike.baidu.com/link?url=YnpcLVjBtFRsfDPnPyssDka7q8BkrQbap-0wuhhgZnpYOM7z OKfpsvwl1ehBFoE



第5章 时间数列

学习目标要求

1. 了解时间数列的概念、时间数列的种类和编制原则。

2. 掌握时间数列水平指标、速度指标的计算和分析方法。掌握各指标之间的关系和应用条件。

重点难点提示

1. 时期数列与时点数列的区别。
2. 时间数列的平均发展水平、增长量和平均增长量等指标的计算及分析。
3. 时间数列的发展速度、增长速度等指标的计算及分析。

教学引导案例

2016年上半年国民经济运行总体平稳、稳中有进

2016年上半年，面对错综复杂的国内外形势和持续较大的经济下行压力，在党中央、国务院的坚强领导下，各地区、各部门认真贯彻落实“创新、协调、绿色、开放、共享”五大发展理念，在适度扩大总需求的同时，加快推进供给侧结构性改革，大力推动大众创业、万众创新，国民经济运行总体平稳、稳中有进。

初步核算，上半年国内生产总值340 637亿元，按可比价格计算，同比增长6.7%。分季度看，一季度同比增长6.7%，二季度增长6.7%。分产业看，第一产业增加值22 097亿元，同比增长3.1%；第二产业增加值134 250亿元，增长6.1%；第三产业增加值184 290亿元，增长7.5%。从环比看，二季度国内生产总值增长1.8%。

一、农业生产基本稳定，夏粮产量略有减少

全国夏粮总产量13 926万吨，比上年减少162万吨，下降1.2%。夏粮总产量略低于最高年份2015年，为历史第二高产年。上半年，猪牛羊禽肉产量3 853万吨，同比下降1.3%，其中猪肉产量2 473万吨，下降3.9%。生猪存栏40 203万头，同比减少3.7%；生猪出栏31 959万头，同比减少4.4%。

二、工业生产稳中略升，盈利状况有所改善

上半年，全国规模以上工业增加值按可比价格计算同比增长 6.0%，增速比一季度加快 0.2%。分经济类型看，国有控股企业增加值同比下降 0.2%，集体企业增长 2.6%，股份制企业增长 7.2%，外商及港澳台商投资企业增长 3.2%。分三大门类看，采矿业增加值同比增长 0.1%，制造业增长 6.9%，电力、热力、燃气及水生产和供应业增长 2.6%。工业结构继续优化。上半年，高技术产业和装备制造业增加值同比分别增长 10.2%和 8.1%，增速分别比规模以上工业快 4.2%和 2.1%，占规模以上工业增加值比重分别为 12.1%和 32.6%，比上年同期提高 0.7%和 1.2%。6 月份，规模以上工业增加值同比增长 6.2%，环比增长 0.47%。上半年，规模以上工业企业产销率达到 97.3%。规模以上工业企业实现出口交货值 55 117 亿元，同比下降 0.7%。

1～5 月份，全国规模以上工业企业实现利润总额 23 816 亿元，同比增长 6.4%（上年同期为下降 0.8%）。规模以上工业企业每百元主营业务收入中的成本为 85.73 元，主营业务收入利润率为 5.59%。

三、固定资产投资增速回落，商品房待售面积减少

上半年，全国固定资产投资（不含农户）258 360 亿元，同比名义增长 9.0%（扣除价格因素实际增长 11.0%），增速比一季度回落 1.7%。其中，国有控股投资 91 089 亿元，增长 23.5%；民间投资 158 797 亿元，增长 2.8%，占全部投资的比重为 61.5%。分产业看，第一产业投资 7 460 亿元，同比增长 21.1%；第二产业投资 101 702 亿元，增长 4.4%；第三产业投资 149 198 亿元，增长 11.7%，其中，基础设施投资（不含电力、热力、燃气及水生产和供应业）49 085 亿元，增长 20.9%。从到位资金情况看，上半年到位资金 282 443 亿元，同比增长 8.0%。其中，国家预算资金增长 21.8%，国内贷款增长 12.4%，自筹资金增长 1.4%，利用外资下降 7.2%。上半年，新开工项目计划总投资 240 202 亿元，同比增长 25.1%。从环比看，6 月份固定资产投资（不含农户）增长 0.45%。

上半年，全国房地产开发投资 46 631 亿元，同比名义增长 6.1%（扣除价格因素实际增长 8.0%），增速比一季度回落 0.1%。其中，住宅投资增长 5.6%。房屋新开工面积 77 537 万平方米，同比增长 14.9%，其中住宅新开工面积增长 14.0%。全国商品房销售面积 64 302 万平方米，同比增长 27.9%，其中住宅销售面积增长 28.6%。全国商品房销售额 48 682 亿元，同比增长 42.1%，其中住宅销售额增长 44.4%。房地产开发企业土地购置面积 9502 万平方米，同比下降 3.0%。6 月末，全国商品房待售面积 71 416 万平方米，比 3 月末减少 2 100 万平方米，比 5 月末减少 753 万平方米。上半年，房地产开发企业到位资金 68 135 亿元，同比增长 15.6%。

……

七、居民收入稳定增长，农民工收入继续增加

上半年，全国居民人均可支配收入 11 886 元，同比名义增长 8.7%，扣除价格因素实际增长 6.5%。按常住地分，城镇居民人均可支配收入 16 957 元，同比名义增长 8.0%，扣除价格因素实际增长 5.8%；农村居民人均可支配收入 6 050 元，同比名义增长 8.9%，扣除价格因素实际增长 6.7%。城乡居民人均收入倍差为 2.80，比上年同期缩小 0.03。全国居民人均可支配收入中位数为 10 505 元，同比名义增长 8.3%。二季度末，农村外出务工劳动力总量 17 509 万人，同比增加 73 万人，增长 0.4%。外出务工劳动力月均收入 3 202 元，同比增长 6.7%。

八、经济结构持续优化，供给侧结构性改革取得进展

产业结构继续优化。上半年，第三产业增加值占国内生产总值的比重为 54.1%，比上年

同期提高 1.8%，高于第二产业 14.7%。需求结构进一步改善。上半年，最终消费支出对国内生产总值增长的贡献率为 73.4%，比上年同期提高 13.2%。中西部地区后发优势继续发挥。上半年，中、西部地区规模以上工业增加值同比分别增长 7.3%和 7.2%，快于东部地区 0.9%和 0.8%；中、西部地区固定资产投资（不含农户）同比分别增长 12.8%和 13.5%，快于东部地区 1.8%和 2.5%。节能降耗继续取得新进展。上半年，单位国内生产总值能耗同比下降 5.2%。

“三去一降一补”成效初显。上半年，原煤、粗钢产量同比分别下降 9.7%和 1.1%。工业企业和商品房库存出现积极变化。5 月末，规模以上工业企业产成品存货同比下降 1.1%。3～6 月，商品房待售面积连续 4 个月减少。工业企业资产负债率及成本均有所下降。5 月末，规模以上工业企业资产负债率为 56.8%，比上年同期下降 0.5%；1～5 月份，规模以上工业企业每百元主营业务收入中的成本比上年同期减少 0.22 元。薄弱领域投资增长较快。上半年，水利环境和公共设施管理业、信息传输软件和信息技术服务业投资同比分别增长 26.7%和 22.5%，分别快于全部投资 17.7%和 13.5%。

总的来看，今年以来，在改革创新深入推进和宏观政策效应不断释放的共同作用下，国民经济保持了总体平稳、稳中有进、稳中有好的发展态势，为完成全年经济目标奠定了较好基础。但也要看到，国内外环境依然复杂严峻，经济下行压力仍然较大。下一步，要坚决贯彻落实党中央、国务院各项决策部署，坚定信心，主动作为，坚持稳中求进工作总基调，适度扩大总需求，坚定不移推进供给侧结构性改革，培育新的经济结构，强化新的发展动能，促进国民经济持续健康发展。

资料来源：http://www.stats.gov.cn/tjsj/zxfb/201607/t20160715_1377652.html

思考：

1. 通过对数据的分析，你对我国 2016 年上半年的经济形势有何判断？
2. 什么是供给侧改革，你的理解是什么？

第一节　时间数列概述

客观物质世界中的一切事物都处在不断发展变化之中。社会经济现象作为客观物质世界的一个重要组成部分，它的规模、结构及现象间的相互联系，随着时间的推移，也都在不断地发展变化着。统计作为认识社会的重要武器，不仅要从现象的相互联系之中进行静态研究，还要从它们的发展变化过程中进行动态研究。

任何事物发展变化都可以从三个方面来反映：过去、现在和未来。通过前面各个章节，我们基本掌握了反映现象数量变化的现状，如用综合指标从静态反映社会经济现象总规模、总水平，用平均发展水平、相对指标反映内部结构、比例、增长状况等。如果想从动态上分析和研究这些事物发展的历史或运行轨迹、发展变化及其规律性，就必须依据时间数列来研究。

时间数列分析是一种广泛使用的统计分析方法，它主要用于描述和探索现象随时间发展变化的数量规律性，是我们研究事物发展状况及预测未来发展趋势的基础和前提条件，所以在社会经济现象动态分析中有着十分重要的作用。

一、时间数列的概念和作用

（一）时间数列的概念

时间数列又称动态数列，它是将社会经济现象某种统计指标的数值按照时间的先后顺序加以排列而形成的统计数列，一般用 x_t 或 y_t 表示。例如将历年来我国国民生产总值加以排列，或者将某企业某年各月的产品产量、职工人数、劳动生产率加以排列，都可以形成时间数列。

时间数列是统计数列的一种，它由以下两个基本要素构成。

（1）现象发展水平所属的时间。可以是某一段时期，如一年、一个月、一天，也可以是某一时点，如某年年底、某月月底。

（2）反映现象发展水平的统计指标数值。

由此可知，时间数列具有两个特点：一是反映客观现象的指标概念相对稳定；二是指标数值随着时间的变化而不断变化。

（二）时间数列的作用

（1）通过时间数列的编制和分析，可以从数量方面研究社会现象的发展过程和规律，并预测未来，为编制计划、制定政策提供依据。

（2）利用有关联的时间数值的对比，分析现象之间发展变化的依存关系和不同空间的发展水平。

二、时间数列的种类

时间数列按其排列指标性质的不同，可以分为总量指标时间数列、相对指标时间数列和平均指标时间数列。其中，总量指标时间数列是基本数列，也称原始数列；相对指标时间数列和平均指标时间数列则是派生数列，因为它们都是根据总量指标时间数列加工计算得到的。

（一）总量指标时间数列

总量指标时间数列又称绝对数时间数列，总量指标时间数列中排列的各项都是绝对数指标，它反映社会经济现象的绝对水平或规模的发展变化过程。总量指标时间数列按其反映的时间状况不同又可分为时期数列和时点数列。

1. 时期数列

时期数列是由时期总量指标编制的时间数列。数列中每项指标数值均表明某种社会经济现象在一定时期内发展过程的总量，如表 5-1 所示。

表 5-1　2007—2014 年我国交通运输、仓储和邮政业增加值

年　份	2007	2008	2009	2010	2011	2012	2013	2014
增加值（亿元）	14 601.0	16 362.5	16 727.1	19 132.2	22 432.8	24 660.0	26 036.3	28 750.0

资料来源：《中国统计年鉴 2012》和《中国统计年鉴 2015》，表中数据按当年价格计算。

表 5-1 中的时间数列，每项指标数值所反映的都是某一段时间内（一年）的交通运输、仓储和邮政业增加值资料，因此为时期数列。由表 5-1 可知，2007—2014 年，我国交通运输、仓储和邮政业增加值不断增加，按当年价格计算，七年期间翻了一番。时期数列中的各项指标所涉及的时间长度称为“时期”，指标数值的大小与时期的长短有直接关系。时期越长，数值越大；时期越短，数值越小。因此，时期数列具有可加性，相加的结果反映现象在更长时间内的发展总量。

时期数列的主要特点有：

（1）时期数列中的各个指标可以相加，各个指标相加具有一定得社会经济意义，反映现象在更长一段时期发展过程的总量；

（2）时期数列中各个指标数值大小与时间长短有直接关系，一般来说，指标所属时间越长，指标数值越大，反之则越小；

（3）时期数列中的各个指标通常是通过连续登记取得的。

2. 时点数列

时点数列是用时点总量指标编制的时间数列。数列中的每项指标数值都反映现象在某一时刻或时点上所达到的总量水平，如表 5-2 所示。

表 5-2　2007—2014 年我国交通运输、仓储和邮政业城镇单位就业人员数（年底数）

年　份	2007	2008	2009	2010	2011	2012	2013	2014
城镇单位就业人员数（万人）	623.1	627.3	634.4	631.1	662.8	667.5	846.2	861.4

资料来源：《中国统计年鉴 2015》。

由表 5-2 可知，2007—2014 年，我国交通运输、仓储和邮政业城镇单位就业人员数有不断增长的趋势，七年间累计增加 238.3 万人，年均增加 34.04 万人。在时点数列中，两个相邻时点之间的时间间隔称为“时点间隔”。它可以是相等的，也可以是不等的。本例的时点间隔是相等的，即均为一年。

设计时点数列时，时点间隔的长短要根据研究对象的特点和研究的任务来决定。一般来说，被研究现象变动较快、较大的，时点间隔应短一些，如银行现金结存余额每天都要结算；反之，对那些变动不大或发展比较缓慢的现象，时点间隔可以长一些，如企业机器设备台数可一年清点一次。

时点数列的主要特点有：

（1）时点数列中的各个指标不能相加，各个指标相加不具有实际的社会经济意义；

（2）时点数列中各个指标数值大小与时间长短没有直接关系；

（3）时点数列中的各个指标通常是通过一定时期登记一次而取得的。

（二）相对指标时间数列

相对指标时间数列是由不同时期的同类相对指标按时间先后顺序排列而成的数列。它反映社会经济现象之间相互联系的发展过程。我国人口出生率就是一个相对指标时间数列，相

对指标时间数列中各项指标数值直接相加没有意义。相对指标时间数列如表 5-3 所示。

表 5-3　2007—2014 年我国交通运输、仓储和邮政业增加值占 GDP 比重

年　份	2007	2008	2009	2010	2011	2012	2013	2014
比重（%）	5.4	5.2	4.8	4.6	4.5	4.4	4.4	4.5

资料来源：《中国统计年鉴 2015》。

由表 5-3 可知，2007—2014 年，我国交通运输、仓储和邮政业增加值占 GDP 比重呈现出先下降后保持稳定的特征，由 2007 年的 5.4%下降至 2011 年的 4.5%后基本稳定，这与服务业增加值占 GDP 比重持续增加的特征是不同的。

（三）平均指标时间数列

平均指标时间数列，是由不同时期的同类平均指标按时间先后顺序排列而成的数列。例如，某大型超市连锁企业的各店铺平均销售额按时间顺序排列资料，就属于平均指标时间数列，如表 5-4 所示。

表 5-4　某大型超市各店铺销售资料

年　份	2012	2013	2014	2015	2016
各店平均销售额（万元）	3 286	6 056	10 687	13 696	14 820

平均指标时间数列能分析社会经济现象一般水平的发展趋势。平均指标时间数列也是由总量指标时间数列对比计算而来的，平均指标时间数列中各项指标数值直接相加没有意义。

总量指标时间数列是基本的时间数列，相对指标时间数列和平均指标时间数列是派生数列。这三种时间数列是相互联系的，分别从不同的角度说明了不同的问题。因此，为了对某一社会经济现象的发展变化过程进行全面分析研究，往往需要将这三种时间数列结合起来运用。

三、时间数列的编制原则

编制时间数列的目的是通过时间数列中各个指标值的对比来研究社会经济现象的发展变化及其规律性，因而保证时间数列中各指标值具有可比性是编制时间数列的基本原则，具体要求如下。

（一）时间的间隔期应长短一致

对于时期数列而言，由于时期数列中指标值的大小与指标所包含的时期长短有直接关系，所以各指标数值包含的时期长短应该一致，否则很难做出判断和比较。

对于时点数列而言，也要求各时点间隔尽可能保持一致，以便更准确地反映现象发展趋势和变化规律。应当指出，对时间可比性的理解不能绝对化。有时为了特定目的也可以把时间长短不同的指标编为时间数列。例如，我国几个重要时期的钢产量资料如表 5-5 所示。

表 5-5　我国几个重要时期的钢产量资料表

时　间	1900—1949	1953—1957	1981—1985	1986—1990	1991—1995
钢产量（万吨）	776	1667	20 304	27 372	42 478

上述时间数列的资料能够很充分地说明新中国成立前经济落后和新中国成立后钢铁工业迅速发展的情况。

（二）总体范围要一致

在时间数列中，各个指标所包括的总体范围前后应该一致。例如，研究某地区工业生产的发展情况时，若该地区的行政区划有了变动，则前后指标数值不能直接对比，必须将资料进行调整，以求总体范围的统一。

（三）计算方法要统一

时间数列各项指标的计算口径、计量单位和计算方法应一致。

例如，要研究企业劳动生产率的变化，产量用实物量还是用价值量，人数用全部职工数还是用生产工人数，前后都要求统一。又如，要把不同时期工业产值进行对比，就应该注意价格水平的变化，采用统一的不变价格表示。不然，价格标准不同，就不能从指标的对比中正确反映工农业产值的实际变化程度。

（四）经济含义要统一

时间数列各项指标所反映的内容应该一致，保证同质性。例如，工业企业里的工资总额，按费用要素分组的工资包括全部职工的工资；而按成本项目分组的工资只包括基本生产工人的工资。如果把这样一些指标数值不加区分就编成时间数列，以反映现象的变动规律，就会导致错误的结论。

第二节　时间数列的水平指标

编制时间数列只是为我们进行动态分析和研究提供了数量依据，而要对现象进行分析和研究，则要通过具体的统计指标来实现。

常用的动态分析指标有：发展水平、平均发展水平、增长量、平均增长量、发展速度、增长速度、平均发展速度和平均增长速度等，前面四种用于现象发展的水平分析，属于水平指标；后面四种用于现象发展的速度分析，属于速度指标。

水平指标是速度指标的基础，速度指标是水平指标进一步加工的结果，是动态分析的继续与深入。

现象发展的水平指标：包括发展水平、平均发展水平、增长量和平均增长量。

一、发展水平

发展水平又称发展量，即时间数列中每一项具体的统计指标数值。它反映社会经济现象在各个不同时期或时点上所达到的规模和水平。它是计算各种动态分析指标的基础。发展水平既可以是总量指标又可以是相对指标或平均指标，如表 5-6 所示。

表 5-6　某企业年末职工人数资料表

年　份	2012	2013	2014	2015	2016
职工人数（人）	526	539	720	823	905

发展水平指标根据在动态分析中所处的位置不同可分为最初水平、最末水平、中间水平等。最初水平是时间数列的第一项指标值，通常用 a_0 表示，即表 5-6 中 2012 年末的职工人数 526 人；最末水平是时间数列中最后一项指标值，通常用 a_n 表示；即表 5-6 中 2016 年末的职工人数 905 人；中间水平是时间数列中除第一项指标值和最后一项指标值之外的其余中间各项指标数值，分别用 $a_1, a_2, \cdots, a_{n-1}$ 表示。

根据发展水平在动态分析中的作用不同，可分为基期水平和报告期水平。

作为比较基准时期的发展水平称为基期水平，一般以数列第一项或报告期的前一期为基期，用 a_0 或 a_{n-1} 表示。

报告期水平是指作为研究时期的发展水平，又称计算期水平，通常用 a_n 表示。

发展水平的概念不是固定不变的，它们会随着研究目的的改变而改变。今年的报告期水平可能是将来的基期水平，这个数列的最末水平可能是另一个数列的最初水平。发展水平在用文字表达变化时，习惯用“增加到”、“增加为”、“降低到”、“降低为”表示。

例如，表 5-6 中某企业职工人数由 2012 年末的 526 人增加到 2016 年末的 905 人。运用时，一定不要把“到”和“为”字漏掉，否则，要说明的社会经济现象指标的意义就要发生变化。

二、平均发展水平

（一）平均发展水平的概念和作用

1. 平均发展水平的概念

平均发展水平又称序时平均数或动态平均数，是将时间数列中各个时间或时点上的发展水平加以平均而得的平均数。它把社会现象在不同时间上的变动差异抽象化，从动态上说明现象在某一时间内的一般水平。

序时平均数和一般平均数（静态平均数）的相同点是研究现象的数量差异抽象化，反映

现象的一般水平。但彼此又有区别，主要体现在以下几方面。

①序时平均数所平均的是研究对象在不同时期上的数量表现，从动态上说明其在某一时期发展的一般水平；而一般平均数是从静态上反映现象的同一时间不同单位的一般水平。

②序时平均数是根据时间数列计算的，它所抽象化的是同一现象在不同时间上所表现出来的差异；而一般平均数是根据变量数列计算的，是将总体各个不同单位在同一时间的数量差异抽象化，用以反映总体在具体历史条件下的一般水平。

2. 序时平均数的作用

序时平均数在统计上具有重要作用：

（1）序时平均数反映现象在一定时间上的一般水平，序时平均数可以消除现象在短时间内波动的影响，便于在各段时间之间进行比较，并观察其发展趋势及其规律；

（2）通过序时平均数可以了解时间数列中某些可比性问题，便于对同一现象在不同时期的变化状况进行比较。

（二）序时平均数的计算方法

序时平均数可以由总量指标时间数列计算，也可以由相对指标时间数列或平均指标时间数列计算。

1. 总量指标时间数列序时平均数的计算

总量指标时间数列有时期数列和时点数列两种，利用它们计算平均发展水平时有不同的方法。

（1）由时期数列计算序时平均数

由于时期数列的各项指标数值可以相加，所以可以采用简单算术平均法计算其平均发展水平。其计算公式为：

$$\bar{a}=\frac{a_1+a_2+\cdots+a_n}{n}=\frac{\sum a}{n}$$

式中：$\bar{a}$ 代表平均发展水平；a 代表各时期发展水平；n 代表时期项数。

例如，计算一年的月平均运输量，可把 12 个月的运输量相加再除以 12。一年的季平均运输量则是四季运输量之和除以 4。

【例 5-1】

2016 年 1 ~ 12 月份我国货物周转量（万吨）资料如表 5-7 所示。

表 5-7　2016 年 1～12 月份我国货物周转量

月　份	货物周转量（万吨）	月　份	货物周转量（万吨）	月　份	货物周转量（万吨）
1	337 393	5	369 034	9	394 983
2	208 730	6	368 298	10	388 500
3	356 940	7	367 370	11	392 533
4	365 752	8	380 584	12	389 054

表 5-7 是根据每月货物周转量资料编制的时期数列，由于各月货物周转量高低不等，因而发展变化趋势不够明显。如果计算出各季的月平均货物周转量，就会明显地反映销售趋势。如：

$$\text{第一季度月平均货物周转量}=\frac{337\,393+208\,730+356\,940}{3}=301\,021.0\text{（万吨）}$$

用同样方法计算出其他各季的月平均货物周转量，如表 5-8 所示。

表 5-8　2016 年各季度我国月平均货物周转量

季　度	第一季度	第二季度	第三季度	第四季度
月平均货物周转量（万吨）	301 021.0	367 694.7	380 979.0	390 029.0

如果要反映 2016 年我国每月的货物周转量情况，只要将 1～12 月份各月的货物周转量直接加总再除以 12 即可，其计算结果为 360 084.8 万吨。

（2）由时点数列计算序时平均数

时点数列是瞬间资料，时点数列中的两个时点之间一般都是有一定间隔的。因此，时点数列一般都是不连续数列。但是，如果时点数列的资料是逐日记录又逐日排列的，这时的时点数列就可以看成连续的时点数列。

时点数列分为连续时点数列和间断时点数列，因此根据掌握的资料情况不同，它们的平均发展水平计算方法是不同的。现分别介绍如下。

①连续、时间间隔相等的时点资料。

如果资料以每日的时间顺序排列，称为间隔相等的连续时点数列，用简单算术平均法计算序时平均数。其计算公式为

$$\bar{a}=\frac{\sum a}{n}$$

式中，a 为各期发展水平；n 为时间数列的项数。

【例 5-2】

根据表 5-9 所示的资料，计算该企业 4 月份平均职工人数。

表 5-9　某企业 4 月份每天职工人数　　单位：人

日　期	职工人数	日　期	职工人数	日　期	职工人数
1	110	11	120	21	124
2	110	12	120	22	126
3	110	13	120	23	126
4	115	14	120	24	126
5	115	15	120	25	126
6	115	16	120	26	118
7	115	17	120	27	118
8	115	18	124	28	118
9	115	19	124	29	118
10	120	20	124	30	118

根据计算公式，该企业 4 月份平均职工人数为

$$\bar{a}=\frac{\sum a}{n}=\frac{110+110+110+\cdots+118}{30}=\frac{3570}{30}=119\text{（万元）}$$

②连续、时间间隔不相等时点资料。

若掌握的时点资料虽是连续的，但没有每日连续登记资料，只有现象发生变化时才登记，则要以时间间隔长度 f 为权数，采用加权算术平均数的方法计算平均发展水平。其计算公式为

$$\bar{a}=\frac{\sum af}{\sum f}$$

式中，a 为各期发展水平；f 为各项指标的时间间隔，其余符号与前面相同。



【例 5-3】

某养殖场 2016 年 6 月份生猪存栏变动记录资料如表 5-10 所示，计算 6 月份生猪日均存栏量。

表 5-10　某养殖场 2016 年 6 月份生猪存栏变动记录表

日期（日）	1	7	15	26
生猪存栏（头）	500	550	520	516

根据上述资料的特点及计算过程整理结果，如表 5-11 所示。

表 5-11　某养殖场 2009 年 6 月份生猪存栏变动计算表

日　期	时间长度（日） f	存栏数（头） a	日头数（头） af
1～7	7	500	3500
8～15	8	550	4400
16～26	11	520	5720
27～30	4	516	2064
合计	30	—	15 684

6 月份平均存栏量为

$$\bar{a}=\frac{\sum af}{\sum f}=\frac{15\,684}{30}=523\text{（头）}$$

③不连续、间隔相等的资料。

在实际工作中，对于时点资料，为了简化统计登记手续，往往每隔一定时间登记一次，时点通常定在月（季、年）初或月（季、年）末，这样就形成了间隔相等的间断时点数列。根据这种数列计算平均发展水平，要假定所研究的现象在两个相邻时点之间的变动是均匀的。因而可将相邻两个时点数值相加再除以 2，求得两个时点之间的简单平均数，然后根据这些

平均数，再用简单算术平均数法计算整个所研究的时间内的现象的平均发展水平。

$$\bar{a}=\frac{\frac{a_1}{2}+a_2+\cdots+a_{n-1}+\frac{a_n}{2}}{n-1}$$

式中，$n-1$为间断数目，它比时点数列的资料少一个。

这种计算方法得出的结果只是一个近似值。



【例 5-4】

根据表 5-12 所示的资料，要求计算某企业 2016 年第一季度职工平均人数。

表 5-12　某企业 2016 年第一季度职工人数资料

日　期	1 月 1 日	2 月 1 日	3 月 1 日	4 月 1 日
职工人数（人）	1030	1010	1200	1180

则该企业第一季度平均职工人数为

$$\bar{a}=\frac{\frac{1}{2}a_1+a_2+\cdots+a_{n-1}+\frac{1}{2}a_n}{n-1}$$

$$=\frac{\frac{1030}{2}+1010+1200+\frac{1180}{2}}{4-1}=\frac{3315}{3}=1105(\text{人})$$

由此可见，在间断时点数列间隔相等的情况下，计算平均发展水平，只要将首末两项时点数值折半，加上中间各项数值，再除以项数减 1 即可，这种方法在统计上被称为“首尾折半法”或“简单序时平均法”。

④不连续、间隔不等的时点资料。

若掌握的时点资料是间隔不等的不连续资料，对于一系列不连续、时间间隔又不相等的时点资料，计算其平均发展水平，则可用时间间隔长度 f_i 为权数加权计算。这种方法被称为“加权序时平均法”。

其计算公式为

$$\bar{a}=\frac{\frac{a_1+a_2}{2}f_1+\frac{a_2+a_3}{2}f_2+\cdots+\frac{a_{n-1}+a_n}{2}f_{n-1}}{\sum_{i=1}^{n-1}f_i}$$

式中，f 代表各时点的间隔长度，其余符号与前面相同。



【例 5-5】

根据表 5-13 所示的资料，计算某公司银行存款余额的全年平均数。

表 5-13　2016 年某公司各月银行存款余额资料　　单位：万元

日　期	1 月 1 日	3 月 1 日	7 月 1 日	10 月 1 日	12 月 31 日
银行存款余额	2902	3040	3010	3100	3202

该公司 2016 年银行存款余额的全年平均数为

$$\bar{a}=\frac{\frac{2902+3040}{2}\times 2+\frac{3040+3010}{2}\times 4+\frac{3010+3100}{2}\times 3+\frac{3100+3202}{2}\times 3}{2+4+3+3}$$

=3055（元）

根据间断时点数列计算序时平均数时，假定研究现象在相邻两个时点之间的变动是均匀的，实际上各种现象的变动并不完全如此。因此，为了使计算的结果能尽量反映实际情况，间断时点数列的间隔不宜过大。

2. 相对指标时间数列序时平均数的计算

相对指标时间数列是由具有互相联系的两个总量指标时间数列对比形成的派生数列，因此要先分别计算出这两个总量指标时间数列的平均发展水平，然后将两者进行对比，求出相对指标时间数列的平均发展水平。其计算公式为

$$\bar{c}=\frac{\bar{a}}{\bar{b}}$$

式中，$\bar{c}$ 代表相对指标时间数列的平均发展水平；$\bar{a}$ 代表作为分子的时间数列的平均发展水平；$\bar{b}$ 代表作为分母的时间数列的平均发展水平。



【例 5-6】

某商业零售企业第一季度各月流动资金周转次数如表 5-14 所示。

表 5-14　某商业零售企业第一季度各月流动资金周转资料

项　目	1 月	2 月	3 月
销售收入 a（万元）	900	1030	1100
流动资金平均占用额 b（万元）	500	660	700
流动资金周转次数 c（次）	1.8	1.56	1.57

试计算该工厂第一季度月平均流动资金周转次数。

表 5-14 中销售收入是时期指标，流动资金占用额是时点指标，而流动资金平均占用额则是月初和月末流动资金占用额的平均数。因此根据所给资料的性质分别求出第一季度各月的平均数，再加以对比，计算流动资金平均周转次数。其具体计算方法如下。

$$\text{该工厂第一季度流动资金平均周转次数}=\frac{\text{第一季度月平均销售收入}}{\text{第一季度流动资金月平均占用额}}$$

即：

$$\overline{c}=\frac{\overline{a}}{\overline{b}}=\frac{\frac{900+1\,030+1100}{3}}{\frac{500+660+700}{3}}=\frac{1\,010}{620}=1.63\text{（次）}$$

3. 平均指标时间数列序时平均数的计算

平均指标时间数列有静态平均数时间数列和动态平均数时间数列之分，它们的序时平均数的计算方法是不同的。

静态平均数是由总体标志总量与总体单位总量对比得到的，也是两个总量指标对比的结果。动态平均数时间数列的序时平均数的计算分两种情况：若时间数列中各指标的时间长度一致，则可用简单算术平均法计算；若时间数列中各指标的时间长短不一，则以时期长度为权数计算加权算术平均数。

三、增长量

（一）增长量的概念

增长量是总体报告期水平与基期水平之差，用以反映现象在一定时期内增加或减少的绝对量。其计算公式为：增长量=报告期发展水平-基期发展水平。

当报告期水平大于基期水平时，增长量为正值，表示现象水平的增加；当报告期水平小于基期水平时，增长量为负值，表示现象水平的下降。有些现象以正增长量为好，如运输量或销售量增加；有些则以负增长量为好，如成本降低。

根据比较基期的不同，增长量可分为逐期增长量和累积增长量两种。

逐期增长量是报告期水平减去前一期水平的差额，说明现象逐期增加的数量。

累积增长量则是报告期水平和某一固定时期水平（通常为最初水平）相减的差额，说明某一时期内的总增长量。

（二）增长量的计算公式

逐期增长量：$a_1-a_0,a_2-a_1,\cdots,a_n-a_{n-1}$。

累积增长量：$a_1-a_0,a_2-a_0,\cdots,a_n-a_0$。

累积增长量与逐期增长量之间有如下数量关系。

（1）累积增长量等于相应各个时期逐期增长量之和，即：

$$(a_1-a_0)+(a_2-a_1)+\cdots+(a_n-a_{n-1})=a_n-a_0$$

（2）相邻两期累积增长量之差等于相应的逐期增长量，即：

$$(a_n-a_0)-(a_{n-1}-a_0)=a_n-a_{n-1}$$



【例 5-7】

某商场销售情况如表 5-15 所示，其中第二列所示的是该商场各年的销售额，其销售额的逐期增长量和累计增长量经计算在第三列、第四列示出。

表 5-15　某商场销售额的增长量计算表　　单位：万元

年　份	销售额	逐期增长量	累计增长量
2010	157	—	—
2011	164	7	7
2012	169	5	12
2013	178	9	21
2014	185	7	28
2015	195	10	38
2016	205	10	48

此外，在日常实际统计工作中，为了消除季节变动带来的影响，也经常计算年距增长量指标，即本期发展水平较上年同期发展水平的增长数量，即：

年距增长量=本期发展水平−去年同期发展水平

四、平均增长量

平均增长量是逐期增长量的序时平均数，说明现象在一定时期内平均每期增长的数量，则其计算公式为

$$\text{平均增长量}=\frac{(a_1-a_0)+(a_2-a_1)+\cdots+(a_n-a_{n-1})}{n}=\frac{a_n-a_0}{n}$$

根据表 5-15 所示的资料，可计算出该商场销售额的年平均增长量为

$$\text{平均增长量}=\frac{7+5+9+7+10+10}{6}=8\text{（万元）}$$

或

$$\text{平均增长量}=\frac{48}{6}=8\text{（万元）}$$

这个数值对于制定第二年的计划很有参考价值，2016 年的销售额（205 万元）加上年平均增长量（8 万元），可作为 2017 年的计划目标，也就是 213 万元。

第三节　时间数列的速度指标

时间数列的速度分析就是将动态数列的指标数值进行对比，经过对比而得到的指标称为

速度指标。时间数列的速度指标包括发展速度、平均发展速度、增长速度和平均增长速度。它们之间联系密切，其中发展速度是最基本的速度指标，是统计中广泛应用的动态分析指标。

一、发展速度

发展速度是说明某种社会经济现象发展程度的相对数，是报告期水平与基期水平对比的比值，表示某一现象在这段时间内发展变化的方向和程度。对比结果以倍数或百分数表示。其计算公式为

$$\text{发展速度}=\frac{\text{报告期水平}}{\text{基期水平}}$$

发展速度大于 1（或 100%）表示上升，小于 1（或 100%）表示下降。

根据对比的基期不同，发展速度可分为环比发展速度和定基发展速度两种。

1. 环比发展速度

环比发展速度是以前一期水平为基期计算的发展速度，反映现象逐期发展变动的程度。它的计算公式为

$$\text{环比发展速度}=\frac{\text{报告期水平}}{\text{前一期水平}}$$

即：

$$\text{环比发展速度}=\frac{a_1}{a_0},\frac{a_2}{a_1},\cdots,\frac{a_n}{a_{n-1}}$$

2. 定基发展速度

定基发展速度是各期水平以某一固定基期水平为基期计算的发展速度，反映现象在一个较长时期内的发展变动程度，因此，定基发展速度又称为总发展速度、总速度。它的计算公式为

$$\text{定基发展速度}=\frac{\text{报告期水平}}{\text{某一固定基期水平}}$$

即：

$$\text{定基发展速度}=\frac{a_1}{a_0},\frac{a_2}{a_0},\cdots,\frac{a_n}{a_0}$$

3. 定基发展速度与环比发展速度之间的数量关系

（1）定基发展速度等于相应的各环比发展速度的连乘积。即：

$$\frac{a_1}{a_0}\times\frac{a_2}{a_1}\times\cdots\times\frac{a_n}{a_{n-1}}=\frac{a_n}{a_0}$$

（2）两个相邻定基发展速度之商等于相应的环比发展速度。即：

$$\frac{a_2}{a_0} \div \frac{a_1}{a_0} = \frac{a_2}{a_1}$$

定基发展速度的基期可根据特定的研究目的来考虑。例如考核五年计划的执行情况时，可选择五年计划的前一年为基期；分析企业的生产经营状况时，可选择历史状态最好的时期作为基期。

二、增长速度

增长速度亦称增长率，它是增长量与基期水平之比，是扣除基数之后的一种发展速度，是说明社会经济现象增长程度的相对数，其计算公式为

$$\text{增长速度} = \frac{\text{增长量}}{\text{基期水平}} = \frac{\text{报告期水平} - \text{基期水平}}{\text{基期水平}} = \text{发展速度} - 1$$

增长速度等于发展速度减 1，当发展速度大于 1 时，增长速度为正值，表示现象增长的程度；当发展速度小于 1 时，增长速度为负值，表示现象减少的程度，“负增长”即此种情况。



小贴士

翻番的含义

在实际工作中，我们经常见到“翻番”一词，它也是速度指标。具体地说，翻一番，指标数值为原来的两倍，即增长一倍，称为一个倍增（增长速度 100%）；翻两番是在原来增加一倍的基础上再增加一倍，即为原来的四倍，实则比原来增加三倍。至于翻更多番的情况，可以想象其速度之大。

根据对比的基期不同，增长速度可分为定基增长速度和环比增长速度。定基增长速度表明现象在较长时间内总的增长速度。环比增长速度表明现象逐期增长程度。其计算公式分别为

$$\text{定基增长速度} = \text{定基发展速度} - 1$$

$$\text{环比增长速度} = \text{环比发展速度} - 1$$



【例 5-8】

根据表 5-16 所示的第二列资料，计算出的该商场销售额的发展速度和增长速度，如表 5-16 第三列和第四列所示。

表 5-16　某商场销售额的发展速度与增长速度计算表

年　份	销售额（万元）	发展速度（%）		增长速度（%）	
		定基	环比	定基	环比
2001	157	100.00	—	—	—
2002	164	104.46	104.46	4.46	4.46
2003	169	107.64	103.05	7.64	3.05
2004	178	113.38	105.33	13.38	5.33
2005	185	117.83	103.93	17.83	3.93
2006	195	124.20	105.41	24.20	5.41
2007	205	130.57	105.13	30.57	5.13

在日常实际统计工作中，对于具有季节性变化的社会经济现象，为了消除季节变动带来的影响，通常计算年距发展速度，用以说明本期发展水平较上年同期发展水平对比达到的发展速度。即：

$$\text{年距发展速度}=\frac{\text{本期发展水平}}{\text{上期同期发展水平}}$$

三、平均发展速度和平均增长速度

平均发展速度和平均增长速度在实际工作中是十分重要的分析指标。平均发展速度和平均增长速度统称为平均速度。平均速度是各个时期环比速度的序时平均数，说明社会经济现象在较长时期内速度变化的平均程度。

（一）平均发展速度

平均发展速度是各期环比发展速度的序时平均数，它说明某社会经济现象在一个较长时期内逐期平均发展变化的程度。由于定基发展速度是各环比发展速度的连乘积，所以平均发展速度不能用一般的算术平均法计算，而要用几何平均法或高次方程法来计算。

1. 几何平均法

几何平均法又称水平法。用这种方法计算平均发展速度就是求各环比发展速度的几何平均数。用几何平均法计算平均发展速度的计算公式为

$$\overline{x}=\sqrt[n]{x_1x_2x_3\cdots x_4}=\sqrt[n]{\prod x}$$

式中，$\overline{x}$ 代表平均发展速度；x 代表各环比发展速度；n 代表环比发展速度的项数。

由时间数列中各环比发展速度的计算公式得出，计算平均发展速度的公式还可以表示为

$$\overline{x}=\sqrt[n]{\frac{a_1}{a_0}\times\frac{a_2}{a_1}\times\frac{a_3}{a_2}\times\cdots\times\frac{a_n}{a_{n-1}}}$$

由于各个环比发展速度连乘积等于最后一年的定基发展速度，所以计算平均发展速度的公式也可以表示为

$$\bar{x}=\sqrt[n]{\frac{a_n}{a_0}}$$

由此可得：

$$a_n=a_0\bar{x}^n$$

上式表达了几何平均法的数学思想，即现象从最初水平出发，每期都按照平均发展速度发展，n 期以后可达到最末水平 a_n。在实际工作中，这个数学思想常常用来推算现象的末期水平。

以上计算平均发展速度的两个公式，可根据提供的具体资料选择应用。如果掌握的资料是各年的环比发展速度，可用前一个公式；如果掌握的资料是最初水平和最末水平，可用后一个公式。



【例 5-9】

根据表 5-17 所示的某企业各年销售收入资料计算平均发展速度。

表 5-17　某企业各年销售收入表　　**单位：万元**

年　份	2011	2012	2013	2014	2015	2016
销售收入	5000	5600	6200	6800	7500	8400

解：该企业 2011 年销售收入为 5000 万元，2016 年为 8400 万元。根据公式计算这一期间的平均发展速度为

$$\bar{x}=\sqrt[n]{\frac{a_n}{a_0}}=\sqrt[5]{\frac{8400}{5000}}=\sqrt[5]{1.68}=110.93\%$$

几何平均法适用于呈递增趋势的社会现象，侧重考察和控制现象的末期水平，计算平均发展速度方法简便。但这种方法忽略了中间各期水平，当中间各期水平波动很大、各环比发展速度差异很大时，用几何平均法计算的平均发展速度就不能确切反映实际发展过程了。

2. 高次方程法

高次方程法又称累计法，高次方程法求平均发展速度的数学思想是：现象从最初水平出发，各期都按照平均发展速度发展，则各期的理论水平之和与各期的实际水平之和相等。即：

$$a_0\bar{x}+a_0\bar{x}^2+a_0\bar{x}^3+\cdots+a_0\bar{x}^n=a_1+a_2+a_3+\cdots+a_n$$

整理后，得：

$$a_0(\bar{x}+\bar{x}^2+\bar{x}^3+\cdots+\bar{x}^n)=\sum a_i$$

$$\bar{x}^n+\bar{x}^{n-1}+\cdots+\bar{x}^2+\bar{x}-\frac{\sum a_i}{a_0}=0$$

解这个方程，求出 $\bar{x}$ 的正根，即为高次方程法求得的平均发展速度。解这样高次方程比较麻烦，在实际工作中，通常使用已经编好的《平均发展速度查对表》中的“累计法查对表”，

根据年限和各年发展水平总和为基础的百分比 $\frac{\Sigma a_i}{a_0}$ 直接查表，求得平均发展速度，或用电子计算器计算。

高次方程法适用于年度间变化不稳定的现象，考虑的是总体各期发展水平的总和，并未考虑到每期水平的分布状况。所以，只要 $\frac{\Sigma a_i}{a_0}$ 的数值确定了，平均发展速度也就随之确定了，这是用高次方程法求平均发展速度时所应注意的问题。

应用几何平均法和高次方程法两种方法求平均发展速度的出发点是不同的，因此，对于同一资料采用两种不同的方法计算，一般情况下结果是不相同的，只有当各期环比发展速度都相等时，这两种方法计算的平均发展速度才会相等。

（二）平均增长速度

平均增长速度是环比增长速度的序时平均数，说明社会经济现象在一个较长时期内逐期平均增长的程度。若计算结果为正值，表明被研究现象在一段时间内逐期递增；若计算结果为负值，则表示其逐期递减。

平均增长速度与平均发展速度有密切的联系，其计算公式为

平均增长速度＝平均发展速度-1

例如，根据表 5-17 中某企业各年销售收入资料计算平均增长速度：

平均增长速度=平均发展速度-100%

=110.93%-100%

=10.93%

在计算和应用平均速度时要注意适当选择基期，并注意所选择的资料应具有同质性。在分析较长历史时期资料时，还应结合分段平均速度来补充总平均速度。

平均发展速度和平均增长速度指标在实际统计工作中应用很广泛。它们可以概括反映国民经济建设中取得的成就，说明各个历史阶段中发展或增长的程度，可用于对不同历史时期、不同国家、不同地区的社会经济现象发展情况进行比较。利用这两个指标可编制长期计划、分析计划完成情况，并将此作为编制计划的重要依据之一，还可以利用平均发展速度预测未来发展水平等。

对于各种有关现象的平均速度指标，可结合起来加以分析，以深入研究有关现象在各个研究时期的平均发展或增长的程度及其差别程度。



小贴士

发展速度与增长速度的关系

发展速度与增长速度是一个问题的两种表述方法，二者有着密切的关系。

（1）发展速度是说明报告期水平发展到基期水平的百分之几，包括了基期水平；而增长速度则说明报告期水平比基期水平增加了百分之几，扣除了基期水平。

（2）发展速度是通过报告期水平与基期水平对比计算的；增长速度是通过报告期水平减

去基期水平后与基期水平对比计算的。

（3）发展速度没有正负值之分，而增长速度则有正负值之分。当增长速度为正值时，表示现象的增长程度；当其为负值时，表示现象的降低程度。

四、增长 1%的绝对值

由于速度指标都是相对数，它会掩盖现象的绝对水平，因此，在经济分析中必须把相对指标与绝对指标结合应用才能全面说明问题。增长 1%的绝对值指标就是把速度指标和发展水平结合运用的一个动态指标。

增长 1%的绝对值，是指在环比增长速度中，报告期水平比前期水平每增长 1%所增长的绝对数量，它等于逐期增长量除以环比增长速度，即前一期水平的 1%。它说明增长速度与绝对量之间的关系，表明研究现象每增长 1%的份额所包含的绝对值。

$$1\%\text{的增长绝对值}=\frac{\text{报告期逐期增长量}}{\text{报告期环比增长速度}}\times 1\%=\frac{\text{基期水平}}{100}$$

本章小结

时间数列是将反映社会经济现象的某一指标在不同时间上的指标数值按照时间的先后顺序排列而成的数列，又称动态数列或时间序列。时间数列的水平指标有发展水平、平均发展水平、增长量和平均增长量。时间数列的速度指标有发展速度、增长速度、平均发展速度和平均增长速度。

复习思考题

1．什么是时间数列？时期数列与时点数列有何区别？

2．时间数列按指标的性质不同可分为哪几类？

3．发展速度与增长速度的关系是怎样的？

4．环比发展速度和定基发展速度的关系是什么？

5．为什么要注意速度指标和水平指标的结合运用？如何结合？



实训基础训练

一、单项选择题

1．已知某企业 7 月、8 月、9 月、10 月的平均职工人数分别为 390 人、398 人、396 人和 358 人，则该企业三季度的平均职工人数应用（　　）计算。

A．首末折半法　　B．简单平均法

C．几何平均法　　D．加权平均法

2．根据时期数列计算序时平均数应采用（　　）。

A．几何平均法　　B．加权算术平均法

C．简单算术平均法　　D．首末折半法

3．说明现象在较长时期发展的总速度的指标是（　　）。

A．环比发展速度　　B．平均发展速度

C．定基发展速度　　D．定基增长速度

4．平均发展速度是（　　）。

A．定基发展速度的算术平均数　　B．环比发展速度的算术平均数

C．环比发展速度连乘积的几何平均数　　D．增长速度加上 100%

5．若各年环比增长速度保持不变，则各年增长量（　　）。

A．逐年增加　　B．逐年减少

C．年年保持不变　　D．无法判断

二、判断题

1．时间数列中，各个环比增长速度的连乘积不等于总的定基增长速度。（　　）

2．发展水平就是动态数列中的每一项具体指标数值，它只能表现为绝对值。（　　）

3．时间数列中，各个环比发展速度的连乘积等于总的定基发展速度。（　　）

4．若逐期增长量每年相等，则其各年的环比发展速度是年下降的。（　　）

5．增长 1%的绝对值表示的是，速度指标增长 1%而增加的水平值。（　　）

三、计算题

1．某企业 2016 年第四季度商品库存情况如表 5-18 所示。

表 5-18　某企业 2016 年第四季度商品库存情况表

日　期	9 月 30 日	10 月 31 日	11 月 30 日	12 月 31 日
商品库存额（万元）	30	38	34	28

试计算该企业 2016 年第四季度的商品平均库存额。

2. 某运输企业第二季度有关资料如表 5-19 所示。

表 5-19　某运输企业第二季度有关资料情况表

月份（月）	4	5	6	7
实际产量（吨）	105	110	115	—
月初工人数（人）	50	50	52	56

试计算：

（1）第二季度平均每月实际产量；

（2）第二季度平均工人数；

（3）第二季度月平均劳动生产率。

3. 已知 2006 年我国国民总收入为 209 407 亿元，若以平均每年增长 9%的速度发展，试计算到 2020 年将达到什么水平？



课外延伸阅读

2016 年 1～5 月份全国规模以上工业企业利润总额同比增长 6.4%

1～5 月份，全国规模以上工业企业实现利润总额 23 816.4 亿元，同比增长 6.4%，增速比 1～4 月份回落 0.1%。

1～5 月份，在规模以上工业企业中，国有控股企业实现利润总额 4 290.2 亿元，同比下降 7.3%；集体企业实现利润总额 177.8 亿元，下降 2%；股份制企业实现利润总额 15 992.9 亿元，增长 7.6%；外商及港澳台商投资企业实现利润总额 6 082.9 亿元，增长 6%；私营企业实现利润总额 8 488.9 亿元，增长 9.4%。

1～5 月份，采矿业实现利润总额 66.6 亿元，同比下降 93.8%；制造业实现利润总额 21 632.6 亿元，增长 12.5%；电力、热力、燃气及水生产和供应业实现利润总额 2 117.2 亿元，增长 1.3%。

1～5 月份，在 41 个工业大类行业中，32 个行业利润总额同比增加，1 个行业持平，8 个行业减少。主要行业利润情况如下：农副食品加工业利润总额同比增长 11.7%，纺织业增长 6.9%，石油加工、炼焦和核燃料加工业增长 3 倍，化学原料和化学制品制造业增长 14.7%，非金属矿物制品业增长 4.1%，黑色金属冶炼和压延加工业增长 74.8%，有色金属冶炼和压延加工业增长 13.4%，通用设备制造业增长 2%，专用设备制造业增长 2.3%，汽车制造业增长 8.9%，电气机械和器材制造业增长 18.4%，计算机、通信和其他电子设备制造业增长 13.5%，煤炭开采和洗选业下降 73.4%，电力、热力生产和供应业下降 0.1%，石油和天然气开采业由同期盈利转为亏损。

1～5 月份，规模以上工业企业实现主营业务收入 426 285.6 亿元，同比增长 2.9%；发生主营业务成本 365 458 亿元，增长 2.7%；主营业务收入利润率为 5.59%。

5 月末，规模以上工业企业资产总计 985 658.7 亿元，同比增长 5.8%；负债合计 559 587.4 亿元，增长 4.9%；所有者权益合计 426 071.3 亿元，增长 7.1%；资产负债率为 56.8%。

5 月末，规模以上工业企业应收账款 114 673.8 亿元，同比增长 8.6%；产成品存货 36 954.7 亿元，下降 1.1%。

1～5 月份，规模以上工业企业每百元主营业务收入中的成本为 85.73 元，每百元资产实现的主营业务收入为 105.3 元，人均主营业务收入为 114.6 万元，产成品存货周转天数为 14.9 天，应收账款平均回收期为 39.1 天。

5 月份，规模以上工业企业实现利润总额 5 372.2 亿元，同比增长 3.7%，增速比 4 月份回落 0.5%（见图 5-1～图 5-3，表 5-20～表 5-21）。

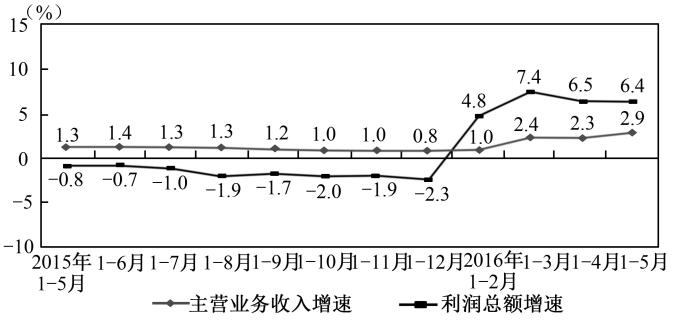


图 5-1　各月累计主营业务收入与利润总额同比增速

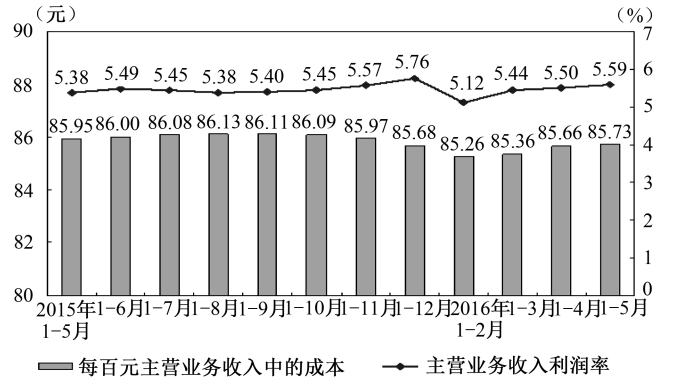


图 5-2　各月累计利润率与每百元主营业务收入中的成本

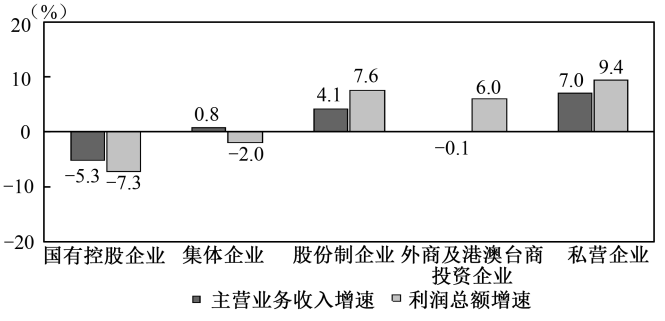


图 5-3　2016 年 1～5 月份经济类型主营业务收入与利润总额同比增速

表 5-20　2016 年 1～5 月份规模以上工业企业主要财务指标

分　组	主营业务收入		利润总额	
	1～5 月（亿元）	同比增长（%）	1～5 月（亿元）	同比增长（%）
总计	426 285.6	2.9	23 816.4	6.4
其中：采矿业	17 769.4	−11.6	66.6	−93.8
制造业	384 132.1	4.0	21 632.6	12.5
电力、热力、燃气及水生产和供应业	24 384.1	−1.6	2117.2	1.3
其中：国有控股企业	88 054.9	−5.3	4290.2	−7.3
其中：集体企业	2619.3	0.8	177.8	−2.0
股份制企业	298 871.4	4.1	15 992.9	7.6
外商及港澳台商投资企业	96 168.0	−0.1	6082.9	6.0
其中：私营企业	149 662.5	7.0	8488.9	9.4
注：经济类型分组之间存在交叉，故各经济类型企业数据之和大于总计。				

表 5-21　2016 年 1～5 月份规模以上工业企业经济效益指标

分　组	主营业务收入利润率（%）	每百元主营业务收入中的成本（元）	每百元资产实现的主营业务收入（元）	人均主营业务收入（万元/人）	资产负债率（%）	产成品存货周转天数（天）	应收账款平均回收期（天）
总计	5.59	85.73	105.3	114.6	56.8	14.9	39.1
其中：采矿业	0.37	86.43	45.6	67.0	62.1	15.3	45.4
制造业	5.63	85.66	125.2	115.7	55.3	15.8	39.7
电力、热力、燃气及水生产和供应业	8.68	86.31	41.3	179.5	61.2	0.6	24.9
其中：国有控股企业	4.87	82.25	55.0	133.6	61.8	18.0	45.5
其中：集体企业	6.79	84.95	128.3	106.4	62.0	15.9	24.4
股份制企业	5.35	85.90	104.5	115.9	57.8	15.4	36.4
外商及港澳台商投资企业	6.33	85.06	118.0	106.9	54.1	15.3	53.4
其中：私营企业	5.67	87.57	173.5	112.9	52.3	12.2	27.4

资料来源：中华人民共和国国家统计局网站 http://www.stats.gov.cn/tjsj/zxfb/201606/t20160627_1371103.html.

第6章 统计指数

学习目标要求

1. 了解统计指数的概念、作用、分类，熟悉和掌握通过综合指数法和平均数指数法编制总指数的原理、计算过程和实际意义。

2. 掌握利用指数体系对复杂现象的变动进行因素分析的方法，并能够实际运用。

重点难点提示

1. 数量指标指数和质量指标指数的编制方法。

2. 加权算术平均数指数和加权调和平均数指数的计算。

教学引导案例

2016年7月中国物流业景气指数为54.8%

中国物流与采购联合会发布的2016年7月份中国物流业景气指数为54.8%，较上月回落0.7%；中国仓储指数为53.7%，较上月回升3.1%；中国公路物流运价指数为101.5点，比上月小幅回升0.2%。

中国物流信息中心副主任何辉认为：7月份，受高温多雨等气候因素影响，中国物流业景气指数较上月略有回落，但仍位于景气区间，显示出经济活动仍较旺盛。中国仓储指数较上月回升，连续五个月位于扩张区间,除了需求与季节因素外，显示出在去库存、降成本背景下，生产企业不断提升第三方专业化库存管理应用水平。从中国公路运价周指数的情况看，随着高温多雨等因素影响的加深，公路物流运价波动中小幅回升，但仍低于上年同期水平。物流行业仍处于较为困难的经营状态中。

两项库存指数双升。7月份，平均库存量指数为55.3%，较上月回升2%；库存周转次数指数为55.9%，较上月回升3.8%。两项库存指数双升，且均位于景气区间，显示出物流业务活动依然活跃。

新订单指数回落。7月份，新订单指数为54.8%，较上月回落2%。虽有所回落，但仍位于较高景气区间，显示出物流需求增势将趋稳。

资金周转率指数回升。7 月份，资金周转率指数回升 1.1%～51.8%。显示出资金状况趋于改善，对受洪涝灾害影响较大的企业灾后重建工作有一定的促进作用。

从业人员指数持平。7 月份，从业人员指数为 49.1%，与上月持平，呈较明显的季节性变化特征，从历史数据看，该指数在未来两个月将逐步回升。

固定资产投资完成额指数回落。7 月份，固定资产投资完成额指数回落了 5.7 个百分点，降至 50.8%。主要是受高温多雨等因素影响，东部和中部地区基建项目进展速度放缓，部分停滞。

从后期走势看，新订单指数回落 2 个百分点，降至 54.8%；业务活动预期指数为 58.6%，保持较高水平。预示着物流行业生产活动将逐步企稳。

资料来源：中国物流信息中心，
http://www.clic.org.cn/wljqzs/271437.jhtml。

思考：

1. 物流业景气指数能够反映我国物流业哪些问题？
2. 物流企业资金周转率等指数的回升意味着什么？

第一节　统计指数概述

一、统计指数的含义

统计指数分析方法是一种最古老、最重要且在对现象的总体数量研究中应用非常广泛的统计方法。指数的概念是从对物价变动的研究中产生的，并有一个逐渐扩展的过程。其渊源可以追溯到 1650 年英国人沃恒（R.Voughan）所创编的物价指数上，迄今已有 300 多年的历史。最初的物价指数大多由个人为研究目的而编制，后来各国政府为了经济管理的需要，也陆续开始编制物价指数。

随着社会实践的需要，指数运用范围逐渐扩大为反映各种经济现象变化，指数种类越来越多，其中最常用的有各种价格指数（零售物价指数、生活费用价格指数、股价指数等）、成本指数、工资指数、产品产量指数、劳动生产率指数等。

统计指数是一种常用且重要的统计指标，运用指数可以对很多社会经济问题进行分析。例如，通过消费价格指数的变化可以对通货膨胀现象进行分析；通过生产指数可以反映经济增长的实际水平；通过运输费用指数可以反映物流成本的变化；通过股价指数可以显示股市行情；通过购买力平价指数可以进行经济水平的国际对比等。在经济分析的各个领域，指数方法都获得了广泛的应用，故统计指数常常被称为“经济指数”。

指数作为一种对比性的统计指标具有相对数的形式，通常表现为百分数。若把作为对比基准的水平（基数）视为 100，则所要考察的现象水平相当于基数的多少。

就不同的总体和不同的研究目的而言，对指数有不同的理解，一般来说，指数的含义有广义和狭义之分。

（一）广义指数

就广义指数本身，也有两种不同的理解。一种理解是：凡是相对数都是指数。另一种理解是：凡是反映同类现象数量变动的相对数都是指数。从中可知，指数从本质上来说就是一种相对数。广义指数的两种理解的区别只在于相对数的具体形式和种类。

（二）狭义指数

狭义指数是指反映复杂总体数量变动的相对数。例如，反映货运价格变动及货物运输量变动的运价指数和运输量指数就属于狭义指数的范畴。由于狭义指数是从总的方面来反映复杂总体综合变动的，故又称之为总指数。

狭义指数的相对形式有多种。反映复杂总体在时间上变动情况的是动态指数；反映复杂总体在不同空间上变动情况的是区域指数；而将同一复杂总体的实际和计划对比所得的相对数是计划完成指数。区域指数、计划完成指数与动态指数的计算原理是一样的。

除非特别说明，一般说的指数都是动态指数。本章着重讨论的也是狭义指数中动态指数的编制理论和方法。

二、统计指数的作用

1. 综合反映事物的变动方向和程度

综合反映事物的变动方向和程度是指数的主要作用。无论是综合指数还是平均数指数，其计算结果一般都是用百分比表示的相对数。这个百分比大于或小于 100%，表示事物变动的方向。大于 100%，说明事物变动的方向是上升；小于 100%，说明事物变动的方向是下降。比 100%大多少或小多少，则说明了事物变动的程度。

2. 分析影响各个因素的方向和程度

用来分析受多因素影响的现象总变动中各个因素的影响方向和影响程度。许多社会经济现象都是复杂现象，其变动要受多种因素影响。通过编制各因素指数，可以分析其对总变动指数影响的方向和影响程度。例如分别编制销售量指数和价格指数，分析它们对销售额的影响方向和影响程度；分别编制产量指数和单位产品成本指数，分析它们对总成本的影响方向和影响程度。

3. 研究事物在长时间内的变动趋势

连续编制的动态指数数列，可以用来进行长时期的总体数量特征发展趋势分析。不仅如此，它还可以把相互有联系指标的指数数列加以比较分析。

例如，工农业产品综合比价指数数列，就是从农产品收购价格指数和工业品零售价格指数两个数列的联系中进行分析，以观察在长时间内工农产品比价的发展趋势。

三、统计指数的种类

（一）按指数所要说明的对象范围不同划分

按指数所要说明的对象是总体还是其构成单位的数量变动，可以将指数分为总指数和个体指数。

1. 个体指数

个体指数是反映构成总体的个体数量变动的相对数。个体指数一般用 K 表示。例如，反映某类货物仓储量变动的相对数就是个体指数。

某类货物仓储量指数为

$$K_q = q_1 / q_0$$

个体价格指数为

$$K_p = p_1 / p_0$$

式中，q 代表“量”；p 代表“价格”；下标 1 和 0 分别代表报告期和基期。

2. 总指数

总指数是反映复杂总体综合变动方向和变动程度的相对数。例如，反映所有货物仓储量变动的仓储量指数、反映所有货物仓储价格综合变动程度的仓储价格指数均是总指数的范畴。总指数一般用符号 $\overline{K}$ 表示。

有时，对复杂总体还要进行分类，计算各组（类）的数量变动。因此在总指数与个体指数之间，还存在一种组（类）指数。组指数实质上就是小范围的总指数。

例如，在编制零售商品物价指数时，就需要先将零售商品进行分类，在分类的基础上，从每类中选代表规格品，然后由各类代表规格品的个体指数推算类指数，再由各类指数推算零售商品物价总指数。可见类指数在个体指数与总指数之间起“桥梁”作用。

（二）按指数所反映的指标的性质不同划分

按指数所反映的指标的性质不同，可以将指数分为数量指标指数和质量指标指数。

1. 数量指标指数

数量指标指数也叫数量指数，是指说明总体或个体在规模、水平方面变动的相对数。例如，工业产品产量指数、商品销售量指数、仓储量指数、职工人数指数等都是数量指标指数。个体数量指数用 K_q 表示，数量总指数则用 $\overline{K}_q$ 表示。

2. 质量指标指数

质量指标指数也称为质量指数，是指说明总体或个体内涵数量变动情况的相对数。例如，劳动生产率指数、平均工资指数、价格指数、单位成本指数都属于质量指标指数。个体质量指数以 K_p 表示，质量总指数以 $\overline{K}_p$ 表示。

（三）按指数反映的时间状况不同划分

按指数反映的时间状况不同，可以分为动态指数和静态指数。

动态指数是指同一总体两个不同时间同类指标对比形成的相对数。“2016 年 5 月，某市物价指数实际上升 5.5%”就属于动态指数。

静态指数是指相同时间不同空间的指标数值对比得到的相对数。例如，甲乙两地同类商品价格对比所形成的相对数就是静态指数。

（四）按所采用的基期不同划分

在指数数列中，按所采用的基期不同，统计指数可以分为定基指数和环比指数。

将一系列性质相同的统计指数按时间顺序排列起来，就形成了一个指数数列。如果数列中指数的基期是固定不变的，则此类指数就是定基指数。

如果基期是可变的，每一期都将前一期的数字作为对比的基数，则是环比指数。例如，将历年的国内生产总值与某一固定年份的国内生产总值对比，所形成的指数就是定基指数；如果逐年对比，则是环比指数。

（五）按其编制方法不同划分

在总指数中，按其编制方法不同，又可分为综合指数和平均数指数。

1. 综合指数

综合指数是指利用复杂总体两个时期可比的现象总量进行对比而得到的相对数，它是计算总指数的基本形式。总指数计算分析的其他方法都是以综合指数的编制原理为依据的。

2. 平均数指数

平均数指数是指利用个体指数或类（组）指数，通过加权平均方法计算而得的相对数。它可以反映复杂总体综合变动程度和变动方向。平均数指数是总指数计算的另一种形式。平均数指数在一定条件下是综合指数的变形，但也可以具有独立意义。

第二节　综 合 指 数

统计研究的对象不在于总体中个体的变动，而在于总体的综合运动。例如，在研究人们的生活水平随工资的提高而提高时，就要同时研究物价的变动，而对物价的研究，不仅要关注某种商品物价的变动，还要关注全社会所有商品物价的综合变动。

总指数反映的是复杂现象的综合变动，而复杂现象总体往往是指使用价值不同、计量单位不一，其数量不能直接相加和对比的多种产品或商品共同构成的总体。要反映这些复杂现象总体的综合变动就必须编制总指数。综合指数就是总指数编制方法中的一种形式。

一、综合指数的概念及特点

综合指数是计算总指数的基本形式和出发点，它是由两个综合的总量指标对比形成的指数。在所研究的总量指标中,包含两个或两个以上的因素，将其中一个或一个以上的因素指标固定下来，仅观察其中一个因素的变动，这样编制出来的总指数就叫作综合指数。

综合指数从编制方法来看，具有以下特点。

（1）先综合后对比。即先解决总体中各个个体由于度量单位不同不能直接加总的问题。为此，需要从经济现象的内在联系出发，确定与研究现象相关联的因素，使它成为同度量因素，从而把不能直接相加的指标过渡到能够相加和比较的指标，然后进行对比。

（2）把总量指标中的同度量因素加以固定，以测定所要研究的因素，即指数化指标的变动程度。

例如，若要观察两个时期商品价值总量中的商品数量的变动，就需要把两个时期各种商品的价格作为同度量因素固定在同一时期，以测定两个时期各种商品总数量的变动。



小贴士

同度量因素

同度量因素亦称同度量系数或权数，是指在计算总指数过程中，通过同度量因素把不能进行直接相加的每种代表品的实物量转化为可以比较的价值量的指标。在不同的使用价值还原为价值的过程中，同度量因素不仅起着媒介作用,还起着权数作用。

（3）分子、分母所研究对象的范围原则上必须一致，反映的是该现象的变动程度。

（4）需要全面资料，综合指数的计算对资料要求较高，需要全面资料。

二、综合指数的编制

综合指数又可分为数量指标综合指数和质量指标综合指数两种,它们的编制原则和方法不同。

（一）数量指标综合指数的编制

我们以计算商品销售量指数为例来说明数量指标综合指数编制的一般方法。



【例 6-1】

假设某商店有三种商品的销售量和价格资料，如表 6-1 所示。

表 6-1 某商店销售量和价格资料表

商品名称	计量单位	价格（元）		销售量	
		基期 p_0	报告期 p_1	基期 q_0	报告期 q_1
甲	件	20	14	50	62.5
乙	千克	10	8	75	90
丙	台	5	5	100	115

现在的目的是计算商品销售量总指数，以反映三种商品销售量的总变动。由于三种商品的度量单位不同，所以三种商品的销售量不能直接相加。因此，要以商品的价格作为同度量因素，使不能直接相加的商品销售量过渡到能够相加的商品销售额，然后用两个时期的商品销售额进行对比来求得商品销售量总指数。

由于销售额的变化包含销售量和价格两个因素的变动，因此需要固定其中的同度量因素（即价格），从而观察销售量的变化。因为作为同度量因素的价格既可以固定在基期，也可以固定在报告期，所以采用不同时期的价格将会得到不同的计算结果，并且具有不同的经济内容。

（1）以基期价格作为同度量因素，销售量总指数的计算可用以下公式表示。

$$\overline{K}_q = \frac{\sum p_0 q_1}{\sum p_0 q_0}$$

式中：$\overline{K}_q$——商品销售量总指数；

$\sum p_0 q_1$——按基期价格计算的报告期销售额；

$\sum p_0 q_0$——基期销售额。

上述公式是由德国学者拉斯贝尔提出的，称为拉斯贝尔数量指标指数公式或拉氏数量指标指数公式。

（2）以报告期价格作为同度量因素，销售量总指数的计算可用以下公式表示。

$$\overline{K}_q = \frac{\sum p_1 q_1}{\sum p_1 q_0}$$

式中：$\sum p_1 q_1$——报告期销售额；

$\sum p_1 q_0$——按报告期价格计算的基期销售额。

上述公式是由德国学者派许提出的，称为派许数量指标指数公式或派氏数量指标指数公式。在计算综合指数时，同度量因素究竟选择哪一个时期是一个十分重要的问题，应根据指数的经济内容来决定，计算商品销售量指数的目的是测定各种商品销售量的总变动，因此，计算时应尽量排除价格变动的影响。

如果根据拉氏数量指标指数公式来计算，即同度量因素固定在基期，这时得到的销售额指标的变动中仅包含销售量这个因素的变动，这实际上意味着按原有价格水平来测定销售量的综合变动。所以，拉氏数量指标指数公式符合计算销售量指数的要求。

如果根据派氏数量指标指数公式来计算，即同度量因素固定在报告期，这时虽然在计算过程中价格固定不变，但由于这样计算得到的销售额指标是按报告期价格计算的，实际上其

中已经包含了价格的变动，这就意味着按调整后的价格来测定销售量的综合变动。所以，根据派氏数量指标指数公式计算的销售量指数与计算该指标的目的不吻合。

因此，在计算数量指标综合指数时，应将同度量因素固定在基期，这是计算数量指标综合指数时选择同度量因素时期的一般原则。

根据表 6-1 中的资料计算，结果如表 6-2 所示。

表 6-2　某商店销售量和价格资料表

商品名称	计量单位	价格（元）		销售量		p_0q_0	p_0q_1	p_1q_1
		基期 p_0	报告期 p_1	基期 q_0	报告期 q_1			
甲	件	20	14	50	62.5	1 000	1 250	875
乙	千克	10	8	75	90	750	900	720
丙	台	5	5	100	115	500	575	575
合计	—	—	—	—	—	2 250	2 725	2 170

根据拉氏数量指标指数公式计算得：

$$\overline{K}_q=\frac{\sum p_0q_1}{\sum p_0q_0}=\frac{1250+900+575}{1000+750+500}=\frac{2725}{2250}=121.11\%$$

$$\sum p_0q_1-\sum p_0q_0=2725-2250=475\ （元）$$

计算结果表明：三种商品销售量报告期比基期增加了 21.11%，由于商品销售量增加而使商品销售额增加了 475 元。

（二）质量指标综合指数的编制

我们仍以表 6-1 中的资料，通过计算商品零售价格指数来说明质量指标综合指数的一般编制方法。

由于三种商品的零售价格不能直接相加，因此需将商品的销售量作为同度量因素，把不能直接相加的商品零售价格过渡为能够相加的商品销售额，然后由两个时期的商品销售额进行对比求得商品零售价格指数。与计算商品销售量指数相似，计算商品零售价格指数也应将同度量因素固定在同一个时期。

（1）以基期销售量作为同度量因素，商品零售价格指数的计算可用如下公式表示。

$$\overline{K}_p=\frac{\sum p_1q_0}{\sum p_0q_0}$$

式中：$\overline{K}_p$——价格总指数。

上述公式又称为拉斯贝尔质量指标综合指数公式或拉氏质量指标综合指数公式。

（2）以报告期销售量作为同度量因素，商品零售价格指数的计算可用如下公式表示。

$$\overline{K}_p=\frac{\sum p_1q_1}{\sum p_0q_1}$$

上述公式又称为派许质量指标综合指数公式或派氏质量指标综合指数公式。

同计算商品销售量指数一样,在计算商品零售价格指数时也存在选择同度量因素时期的问题。我们计算商品零售价格指数的目的是测定商品价格的波动情况，以说明市场物价变动对人民生活的影响程度。

如果根据拉氏质量指标综合指数公式来计算，即同度量因素固定在基期,其计算结果表明：按过去的商品销售量、商品零售价格的变动程度，公式的分子与分母之差额说明由于物价的变动，居民按过去的购买量购买这三种商品将多支出或少支出的金额，这显然是没有多少现实意义的。

如果根据派氏质量指标综合指数公式来计算，即同度量因素固定在报告期，其计算结果表明：按目前的商品销售量、商品零售价格的变动程度，公式的分子与分母之差额说明由于物价的变动，居民按目前的购买量购买这三种商品将多支出或少支出的金额，这比较符合商品零售价格指数的意义。

因此，在计算质量指标综合指数时，应将同度量因素固定在报告期，这是计算质量指标综合指数时选择同度量因素时期的一般原则。

根据表 6-2 中的资料和派氏质量指标综合指数公式计算商品零售价格指数为

$$\bar{K}_p=\frac{\sum p_1q_1}{\sum p_0q_1}=\frac{875+720+575}{1250+900+575}=\frac{2170}{2725}=79.63\%$$

$$\sum p_1q_1-\sum p_0q_1=2170-2725=555\text{（元）}$$

计算结果表明：三种商品价格报告期比基期下降了 20.37%，由于价格下降，使居民购买这三种商品少支出 555 元。

综上所述，在计算综合指数时，同度量因素时期选择的一般原则是：数量指标综合指数的同度量因素时期固定在基期，质量指标综合指数的同度量因素时期固定在报告期。但也不是一成不变的，在实际应用中，还要根据研究目的来确定同度量因素的时期。

目前，常用不变价格作为同度量因素，即不论计算哪个时期的总指数，都采用同一个特定时期的价格作为同度量因素，其计算公式如下：

$$\bar{K}_q=\frac{\sum p_nq_1}{\sum p_nq_0}$$

式中，p_n 代表不变价格。

第三节　平 均 指 数

一、平均数指数的概念

平均数指数是计算总指数的另一种重要形式，它是从构成复杂社会经济现象的各种因素

的个体指数出发，通过对个体指数进行加权平均而得到的总指数。我们知道，在计算和编制综合指数时需要采用全面资料，如果基期或报告期的资料不全，就不能用综合指数进行计算，这时就必须采用平均数指数进行计算。

依据“先对比，后平均”的形式编制平均指数的基本程序是：首先通过对比计算出个体指数，然后将个体指数进行加权平均求得总指数。

平均指数与综合指数的联系在于：在特定的权数下，两者之间存在变形关系。但是作为一种独立指数形式的平均指数，不只是作为综合指数的变形使用，其本身也具有独立的应用价值。

二、平均数指数的编制

最常用的平均数指数有加权算术平均数指数和加权调和平均数指数两种形式。

（一）加权算术平均数指数的编制

根据基期同度量因素进行计算的综合指数都可以变形为加权算术平均数指数，所以拉氏综合指数都可变形为加权算术平均数指数。

设 $k=\frac{q_1}{q_0}$ 代表个体销售量指数，则拉氏销售量综合指数：

$$\overline{K}_q=\frac{\sum p_0q_1}{\sum p_0q_0}=\frac{\sum \frac{q_1}{q_0}p_0q_0}{\sum p_0q_0}=\frac{\sum kp_0q_0}{\sum p_0q_0}$$

由上述公式可知：$\overline{K}_q$ 是 k 以 p_0q_0 为权数而求得的加权算术平均数。

设 $k=\frac{p_1}{p_0}$ 代表个体价格指数，则拉氏商品价格综合指数：

$$\overline{K}_p=\frac{\sum p_0q_1}{\sum p_0q_0}=\frac{\sum \frac{q_1}{q_0}p_0q_0}{\sum p_0q_0}=\frac{\sum kp_0q_0}{\sum p_0q_0}$$

由上述公式可知：$\overline{K}_p$ 是 k 以 p_0q_0 为权数而求得的加权算术平均数。

假设表 6-1 中只有销售量和基期销售额资料，这时就不能直接采用综合指数公式进行计算，只能用平均数指数公式来计算总指数。计算表如表 6-3 所示。

表 6-3　某商店销售量和个体指数资料表

商品名称	计量单位	销售量		个体指数 $k=q_1/q_0$	基期销售额 p_0q_0	kp_0q_0
		基期 q_0	报告期 q_1			
甲	件	50	62.5	1.25	1000	1250
乙	千克	75	90	1.2	750	900
丙	台	100	115	1.15	500	575
合计	—	—	—	—	2250	2725

根据公式计算得：

$$\overline{K}_q=\frac{\sum kp_0q_0}{\sum p_0q_0}=\frac{2725}{2250}=121.11\%$$

计算结果与采用综合指数公式计算结果一致。这就表明：加权算术平均数指数实际上是数量指标综合指数的变形，两者虽然形式不同，但结果和经济内容是一致的，在以 p_0q_0 为权数的情况下，两者之间可以相互转换。

（二）加权调和平均数指数的编制

根据报告期同度量因素进行计算的综合指数都可以变形为加权调和平均数指数，所以派氏综合指数都可变形为加权调和平均数指数。

设 $k=\frac{q_1}{q_0}$ 代表个体销售量指数，则派氏销售量综合指数：

$$\overline{K}_q=\frac{\sum p_1q_1}{\sum p_1q_0}=\frac{\sum p_1q_1}{\sum \frac{q_0}{q_1}p_1q_1}=\frac{\sum p_1q_1}{\sum \frac{p_1q_1}{k}}$$

由上述公式可知：$\overline{K}_q$ 是 k 以 p_1q_1 为权数而求得的加权调和平均数。

设 $k=\frac{p_1}{p_0}$ 代表个体价格指数，则派氏价格综合指数：

$$\overline{K}_p=\frac{\sum p_1q_1}{\sum p_0q_1}=\frac{\sum p_1q_1}{\sum \frac{p_0}{p_1}p_1q_1}=\frac{\sum p_1q_1}{\sum \frac{p_1q_1}{k}}$$

由上述公式可知：$\overline{K}_p$ 是 k 以 p_1q_1 为权数而求得的加权调和平均数。

假设表 6-1 中只有商品零售价格和报告期销售额资料，就不能直接采用综合指数公式进行计算，只能用平均数指数公式来计算总指数。计算表如表 6-4 所示。

表 6-4　某商店销售价格和个体指数资料表

商品名称	计量单位	价格（元）		个体指数 $k=p_1/p_0$	报告期销售额 p_1q_1	$\frac{p_1q_1}{k}$
		基期 p_0	报告期 p_1			
甲	件	20	14	0.7	875	1250
乙	千克	10	8	0.8	720	900
丙	台	5	5	1	575	575
合计	—	—	—	—	2170	2725

根据公式计算得：

$$\overline{K}_p=\frac{\sum p_1q_1}{\sum \frac{p_1q_1}{k}}=\frac{2170}{2725}=79.63\%$$

计算结果与采用综合指数公式计算的结果一致。这就表明：加权调和平均数指数实际上是质量指标综合指数的变形，两者虽然形式不同，但结果和经济内容是一致的，在以 p_1q_1 为

权数的情况下，两者之间可以相互转换。

此外,在实际工作中，常把平均数指数的权数固定，以正常年份数量构成或价格水平作为编制产量指数或价格指数的权数，使用一段时期不变，这种权数就叫固定权数，按这种方法计算的平均数指数就称为固定权数平均数指数。固定权数平均数指数是计算总指数的一种独立的形式，其计算公式表示如下：

$$\text{固定权数算术平均数指数}=\frac{\sum kw}{\sum w}$$

$$\text{固定权数调和平均数指数}=\frac{\sum w}{\sum \frac{w}{k}}$$

式中：k 代表个体指数，w 代表固定权数。

第四节　指数体系和因素分析

一、指数体系的概念及作用

客观现象是错综复杂的，各种因素对它的影响不是孤立的，而是互相联系、互相制约和相互影响的。统计中，除了依据现象内在因素联系编制综合指数以外，还要应用指数体系来分析现象中各个因素的影响程度，这就需要建立指数体系。

由若干个经济上有联系、数量上保持一定关系的指数所组成的整体就称作指数体系。其中由两个不同时期总量对比形成的指数称为价值指数，又称为总量指数，如不同时期的工业总产值、产品总成本、商品销售额等对比形成的价值指数。

社会经济现象的总量通常可以分解为若干个构成因素，如商品销售额是销售量（q）与销售价格（p）的乘积，粮食总产量可以看作播种面积（q）与单位面积产量（p）的乘积等。

指数体系的主要作用表现在两个方面：一是可以进行指数间的互相推算；二是可以测定各因素的变动对总量指标变动的影响,进行因素分析。

二、因素分析

（一）因素分析的概念

因素分析是依据指数体系理论，分析受多种因素影响的社会经济现象总量指标变动中各因素影响的方向和程度的方法。

因素分析的内容主要包括以下两方面。

（1）从相对数方面分析现象总量指标的变动受各因素变动影响的方向和程度。

各因素指数的乘积等于总变动指数，该指数体系可表示为如下公式：

$$\frac{\sum p_1q_1}{\sum p_0q_0}=\frac{\sum p_0q_1}{\sum p_0q_0}\times\frac{\sum p_1q_1}{\sum p_0q_1}$$

例如，价值指数（$\sum pq$）=数量指标指数（$\sum q$）×质量指标指数（$\sum p$）

总成本指数=产量指数×单位成本指数

总产值指数=产量指数×产品价格指数

销售额指数=销售量指数×价格指数

（2）从绝对数方面分析总量指标的变动受各因素变动影响的水平。

各因素指数分子分母差额的总和等于总量指数实际发生的总差额，公式如下：

$$\sum p_1q_1-\sum p_0q_0=(\sum p_0q_1-\sum p_0q_0)+(\sum p_1q_1-\sum p_0q_1)$$

（二）因素分析的应用

总量指标因素分析的方法有总量指标的两因素分析法、总量指标的多因素分析法等。根据需要下面重点介绍总量指标的两因素分析法。

【例 6-2】

某厂生产两种产品资料，如表 6-5 所示。试从绝对数和相对数两方面分析该厂单位成本和产量的变动对总成本变动影响的方向和程度。

表 6-5　某厂产品产量和成本资料表

产品名称	单　位	产量（万件）		单位成本（元）		总成本（万元）		
		基期 q_0	报告期 q_1	基期 z_0	报告期 z_1	z_0q_0	z_1q_1	z_0q_1
甲	米	65	40	8	9	520	360	320
乙	台	50	75	6	4	300	300	450
合计	—	—	—	—	—	820	660	770

解：总成本变动受产量和单位成本两个因素变动的影响，故可依据下列指数体系来分析总成本变动的影响因素：

总成本指数 = 产量指数×单位成本指数

$$\frac{\sum z_1q_1}{\sum z_0q_0}=\frac{\sum z_0q_1}{\sum z_0q_0}\times\frac{\sum z_1q_1}{\sum z_0q_1}$$

$$\sum z_1q_1-\sum z_0q_0=(\sum z_0q_1-\sum z_0q_0)+(\sum z_1q_1-\sum z_0q_1)$$

利用该指数体系进行因素分析的步骤如下。

第一步，计算总成本指数及其绝对数的增减额。

总成本指数及其增减额：

$$总成本指数=\frac{\sum z_1q_1}{\sum z_0q_0}=\frac{660}{820}=80.49\%$$

$$总成本变动额=\sum z_1q_1-\sum z_0q_0=660-820=-160（万元）$$

第二步，计算两个因素（产量和单位成本）指数及其绝对数的增减额。

产量指数及对总成本影响的增减额:

$$\frac{\sum z_0 q_1}{\sum z_0 q_0} = \frac{770}{820} = 93.90\%$$

$$\sum z_0 q_1 - \sum z_0 q_0 = 770 - 820 = -50 \text{（万元）}$$

单位成本指数及对总成本影响的增减额:

$$\frac{\sum z_1 q_1}{\sum z_0 q_1} = \frac{660}{770} = 85.71\%$$

$$\sum z_1 q_1 - \sum z_0 q_1 = 660 - 770 = -110 \text{（万元）}$$

第三步，因素分析内容。

相对数变动：80.49%=93.90%×85.71%

绝对数变动：−160 万元 =（−50 万元）+（−110 万元）

第四步，从相对数和绝对数两个方面进行分析说明。

从相对数来看，两种产品的总成本报告期比基期下降了 19.51%，是产量下降了 6.1%和单位成本下降了 14.29%两个因素共同作用的结果。从绝对数来看，由于两种产品的总成本报告期比基期减少了 160 万元，是产量下降使总成本降低了 50 万元、单位成本降低使总成本降低了 110 万元两因素共同作用的结果。

第五节　指数在统计实践中的应用

统计指数在我国社会经济生活中应用非常广泛，每月、每季度和每年国家统计局都要公布一些统计指数来说明经济运行情况，并用来预测经济走势。常用的统计指数主要有商品零售价格指数（RPI）、居民消费价格指数（CPI）、生产者物价指数（PPI）和股票价格指数等。其中与人们生活关系最为密切的是商品零售价格指数和居民消费价格指数。

一、商品零售价格指数

商品零售价格指数（Retail Price Index，RPI）是全面反映工业、商业、餐饮业和其他零售企业向城乡居民、机关团体出售生活消费品和办公用品价格变动趋势和变动程度的相对数。编制 RPI 的目的在于掌握零售商品价格的变动趋势，为国家宏观调控和国民经济核算提供参考依据。

商品零售价格的调查范围涉及各种类型的工业、商业、餐饮业和其他行业等零售商品的价格，包括食品、饮料烟酒、服装鞋帽、纺织品、家用电器及音像器材、文化办公用品、日用品、体育娱乐用品、交通通信用品、家具、化妆品、金银珠宝、中西药品及医疗保健用品、

书报杂志及电子出版物、燃料、建筑材料及五金电料等 16 个大类。

商品零售价格指数的编制方法如下：

（1）选择恰当的调查地区和调查点；

（2）选择代表商品和代表规格品；

（3）搜集价格资料；

（4）确定权数；

（5）计算平均价格；

（6）编制价格指数。

二、居民消费价格指数

居民消费价格指数简称（Consumer Price Index，CPI）反映一定时期内居民所消费商品及服务项目的价格水平变动趋势和变动程度。居民消费价格指数是用百分数表示的，若高于100%，表示价格水平上升；若低于 100%，表示价格水平下降。CPI 是反映我国通货膨胀（或紧缩）程度的主要指标，在国民经济价格体系中具有极为重要的地位。

通俗地讲，CPI 就是市场上的货物价格增长百分比。一般市场经济国家认为 CPI 增长率在 2%～3%属于可接受范围内，当然还要看其他数据。CPI 过高不是好事，高速经济增长率会拉高 CPI，但物价指数增长速度快过人民平均收入的增长速度就一定不是好事，而一般平均工资的增长速度很难超越 3%～4%。

居民消费价格指数是在对全国 550 个样本市、县的近 3 万个采价点进行价格调查的基础上，根据国际规范的流程和公式算出来的。我国 CPI 的调查内容包括食品、烟酒及用品、衣着等 8 大类。编制 CPI 所用权重是依据全国 12 万户城乡居民家庭调查资料中的消费支出构成确定的。

我国居民消费价格指数是采用固定加权算术平均指数方法编制的。其主要编制过程如下。

（1）将各种居民消费划分为 8 大类，即食品、衣着、家庭设备及用品、医疗保健、交通和通信工具、文教娱乐用品、居住项目及服务项目等，下面再划分为若干个大类、中类和小类。

（2）从以上各类中选定有代表性的商品项目(含服务项目)列入指数，利用有关对比时期的价格资料分别计算个体价格指数。目前选定约 700 个规格品种的商品和服务项目作为我国居民消费价格调查的“商品篮子”，其中不包括投资品和资产项目。

（3）依据有关时期内各种商品的销售额构成确定代表品的比重权数。

（4）按从低到高的顺序，采用固定加权算术平均数指数公式，依次编制各小类、中类、大类的消费价格指数和消费价格总指数。

即：

$$\bar{K}_p = \frac{\sum k_p w}{\sum w}$$

式中，k_p 为商品价格个体指数或各层的类指数，w 为权数。



小贴士

为什么商品房没有纳入 CPI 统计？

商品房没有纳入 CPI 统计有以下原因。

（1）出于国民经济核算的需要。目前世界各国统计机构所采用的“国际惯例”为住宅按新的或现存的有形固定资产来处置。为此，CPI 的统计口径必须与国民经济核算体系中的类相一致，而根据国民经济核算分类，商品房购买属于投资范畴，而 CPI 关注的是消费。

（2）出于测量通货膨胀的需要。测量通货膨胀的指标不止一个，但世界各国大多使用 CPI 来测量。商品和服务价格普遍地、持续地上涨就发生了通货膨胀，而不是资产的价格上涨发生通货膨胀。假设我们把目前商品房价格变动作为消费品价格变动纳入 CPI 的统计中，那么 CPI 会更高。这与现实商品和服务价格情况不符，容易引起误解。

（3）商品房购买与当期消费不同步，购买支出与当期实际住房消费不对等。商品房购买行为实质上是一种在短期内大量金额的集中支付，但商品房要用于今后几十年的消费。也就是说，当期的实际住房消费对应的只是整个住房的一部分（即折旧），而不是整个住房。

（4）事实上 CPI 已经反映了居住消费价格的变动。不将商品房（资产）价格变动纳入 CPI 统计范畴，不代表 CPI 不反映居住类（消费）价格的变化情况。

三、生产者物价指数

生产者物价指数简称 PPI，它是站在生产者的角度来衡量工业企业产品出厂价格变动趋势和变动程度的指数，是反映某一时期生产领域价格变动情况的重要经济指标，也是制定有关经济政策和国民经济核算的重要依据。

根据价格传导规律，PPI 反映生产环节的价格水平，CPI 反映消费环节的价格水平，PPI 对 CPI 有一定的影响。因此，PPI 能够反映生产者获得原材料的价格波动情况，推算预期 CPI，从而估计通胀风险。2009 年 1 月开始，我国开始同时公布 CPI 和 PPI，为估计和预测经济运行情况提供依据。

在我国，PPI 一般指统计局公布的工业品出厂价格指数，目前我国 PPI 的调查产品有 4000 多种，覆盖全部 39 个工业行业大类，涉及调查种类 186 个。其中，能源原材料价格在 PPI 构成中占较大比重。通常情况下，PPI 走高意味着企业出厂价格提高，因此会导致企业盈利增加；但如果下游价格传导不利或市场竞争激烈，走高的 PPI 则意味着众多竞争性领域的企业将面临越来越大的成本压力，从而影响企业盈利，整个经济运行的稳定性也将受到考验。

四、股票价格指数

（一）股票价格指数概念

股票价格指数一般采用与基期比较法，即将选样股票计算期的价格总和与基期的价格总和进行比较，反映各个时期价格水平的变动情况，简称股价指数。指数单位一般用“点”表示，“点”是衡量股票价格起落的尺度，即将基期指数作为 100，每上升或下降一个百分点称为“1 点”。

编制股价指数时选择有代表性的、实力雄厚的上市公司的股票，这些公司的股价变动反映了股市的股价水平，而这些公司的经营业绩又反映了该国家或地区的经济状况，所以股价指数是观察和分析经济的重要参考依据。

（二）选择权数的方法

一般方法是以采样股股票发行量为权数，以求得市价总值，报告期市价总值与基期市价总值之比即为指数。以市价总值股价指数显示采样股整体资产价值的变化，相对地降低了某些股票市价暴涨暴跌对指数的影响。当发生股票分割、配股时，采样股数量相应地逐渐膨胀，权数增大，以致出现失真现象，故以发行量为权重，比较适合于观察平均股价水准变动。

另一种方法是以采样股股票成交量为权数，但该种权数是不固定的，当股价上涨而成交量骤变时，股价指数容易使投资者产生误解。但以成交量为权重，反映了平均每股成交额，可用以测定股票市场的投资者心态。

（三）著名的股价指数

1. 道·琼斯股票价格指数

道·琼斯股票价格指数又称道·琼斯股票价格平均指数，是世界上最有影响、使用最广的股价指数。它以在纽约证券交易所挂牌上市的一部分有代表性的公司股票作为编制对象，由四种股价平均指数构成：

①以 30 家著名的工业公司股票为编制对象的道·琼斯工业股价平均指数；

②以 20 家著名的交通运输业公司股票为编制对象的道·琼斯运输业股价平均指数；

③以 6 家著名的公用事业公司股票为编制对象的道·琼斯公用事业股价平均指数；

④以上述三种股价平均指数所涉及的 56 家公司股票为编制对象的道·琼斯股价综合平均指数。

在四种道·琼斯股价指数中，以道·琼斯工业股价平均指数最为著名，它被大众传媒广泛地报道，并作为道·琼斯指数的代表加以引用。

道·琼斯股票价格指数以 1928 年 10 月 1 日为基期，采用算术平均法进行计算编制而成，当遇到股票的除权除息时，股票指数将发生不连续的现象。1928 年后，道·琼斯股票价格平均数改用新的计算方法，即在计点的股票除权或除息时采用连接技术，以保证股票指数的连续性，从而使股票指数得到了完善，并逐渐推广到全世界。

2. 香港恒生指数

香港恒生指数是香港股票市场上历史最悠久、影响最大的股票价格指数，由香港恒生银

行于 1969 年 11 月 24 日开始发布。恒生股票价格指数将从香港 500 多家上市公司中挑选出来的 33 家有代表性且经济实力雄厚的大公司股票作为成分股，分为四大类，即 4 种金融业股票、6 种公用事业股票、9 种房地产业股票和 14 种其他工商业(包括航空和酒店)股票。

恒生股票价格指数的编制以 1964 年 7 月 31 日为基期，因为这一天香港股市运行正常，成交值均匀，可反映整个香港股市的基本情况，基点确定为 100 点。其计算方法是将 33 种股票按每天的收盘价乘以各自的发行股数作为计算日的市值，再与基期的市值相比较，再乘以 100，就得出当天的股票价格指数。

3. 上证综合指数和深证成份股指数

（1）上证综合指数

上证综合指数是由上海证券交易所编制的综合指数类股票指数，简称上证综指，它以 1990 年 12 月 19 日为基准日，基准日指数定为 100 点，1991 年 7 月 15 日正式开始发布。该股票指数的样本为所有在上海证券交易所挂牌上市的股票，其中新上市的股票在挂牌的第二天纳入股票指数的计算范围，该股票指数的权数为上市公司的总股本。

其公式为：

$$\text{本日股价指数} = \text{本日股票市价总值} / \text{基期股票市价总值} \times 100$$

其中：

$$\text{本日股票市价总值} = \sum \text{本日收盘价} \times \text{发行股数}$$

$$\text{基期股票市价总值} = \sum \text{基期收盘价} \times \text{发行股数}$$

随着上市股票品种的逐渐增加，上海证券交易所在这一综合指数的基础上，从 1992 年 2 月起分别公布 A 股指数和 B 股指数，1993 年 5 月 3 日起正式公布工业、商业、地产业、公用事业和综合五大类分类股价指数，以反映不同行业股票的各自走势。

至此，上证指数已发展成为包括综合股价指数、A 股指数、B 股指数、分类指数在内的股价指数系列。

（2）深证成份股指数

深证成份股指数是反映深圳证券交易所成份股价的指数，简称深证成指，它是深圳证券交易所的主要股指之一。深证成指 1995 年 5 月起开始发布，它按一定标准选出 40 家有代表性的上市公司作为成份股，用成份股的可流通数作为权数，以 1994 年 7 月 20 日为基日，基日指数定为 1000 点，以加权平均法进行编制，其基本公式为：

$$\text{股价指数} = \text{现时成分股总市值} / \text{基期成分股总市值} \times 1000$$

深证成份股指数选取样本时考虑的因素有：

①日期的长短；

②上市规模，按每家公司一段时期内的平均总市值和平均可流通股市值计；

③交易活跃程度，按每家公司一段时期总成交金额计。

确定初步名单后，再结合以下各因素评选出 40 家上市公司作为成份股：

①股票在一段时期内的平均市盈率；

②公司的行业代表性及所属行业的发展前景；

③公司近年的财务状况、盈利记录、发展前景及管理素质等；

④公司的地区、板块代表性等。

为保证指数的代表性，必须视上市公司的变动更换成份股，深圳证券交易所定于每年 1 月、5 月、9 月对成份股的代表性进行考察，讨论是否需要更换。



小贴士

深证 100 指数

深证 100 指数是以深圳市场全部正常交易的股票(包括中小企业板)作为选样范围，选取 100 只 A 股作为样本编制而成的成分股指数，并保证中小企业成分股数量不少于 10 只，属于描述深市多层次市场指数体系核心指数。

该指数于 2003 年 1 月 2 日发布，基点和发布点位为 1000 点。深证 100 指数的功能定位主要就是向市场投资者(特别是机构投资者)提供客观的投资业绩基准和指数化投资标的。深证 100 指数每半年调整一次。

深圳证券信息有限公司推出的深证 100 指数是中国证券市场有史以来第一只由中立机构编制、管理和面向整个证券市场发布的股票指数，体现了指数公正、公开、客观、透明的原则，说明中国证券市场开始向与国际化接轨迈出了历史性的一步。



本章小结

综合指数是实际中应用最多的一种指数形式，其编制方法主要有拉氏指数和派氏指数两种形式。平均数指数则主要作为综合指数的一种变形来使用。平均指数的计算形式有两种：加权算术平均数指数和加权调和平均数指数。指数体系是由总量指数和若干个因素指数构成的数量关系式，指数体系的作用主要表现在两个方面：一是进行因素分析；二是用于指数之间的相互推算。居民消费价格指数、零售价格指数和股票价格指数是几种重要的经济指数，理解它们的编制过程有助于我们对实际经济问题的认识和研究。



复习思考题

1．统计指数的概念和作用有哪些？

2．什么是综合指数？为什么说综合指数是总指数的基本形式？

3．什么是指数体系？怎样利用指数体系进行二因素分析？

4．计算综合指数时，选择同度量因素时期的一般原则是什么？

5．综合指数与平均指数有何区别与联系？



实训基础训练

一、单项选择题

1．总指数有两种计算形式，即（　　）。

A．个体指数和综合指数　　B．综合指数和平均数指数

C．算术平均数指数和调和平均数指数　　D．综合指数和平均指标指数

2．下列指数中属于质量指标指数的是（　　）。

A．产量指数　　B．商品销售量指数　　C．职工人数指数　　D．劳动生产率指数

3．同度量因素的作用有（　　）。

A．平衡作用　　B．比较作用　　C．权数作用　　D．稳定作用

4．对商品销售额的变动情况进行分析时，影响销售额指数的因素包括销售量指数和（　　）。

A．个体指数　　B．销售价格指数　　C．总平均价格指数　　D．销售额指数

5．若价格增长 5%、销售量增长 4%，则销售额增长（　　）。

A．20%　　B. 9%　　C. 9.2%　　D. 8%

6．加权算术平均数指数主要用来编制（　　）指数。

A．数量指标　　B. 质量指标　　C. 相对指标　　D. 平均指标

二、判断题

1．加权调和平均数指数实际上也是综合指数的变形，两者虽然形式不同，但结果和经济内容是一致的。（　　）

2．同度量因素的基本作用是将不能直接相加的数量过渡到能够相加的数量。（　　）

3．某企业某种产品产量提高 6%，出厂价格下降 6%，则其产品产值不变。（　　）

4．数量指标指数是以数量指标作为同度量因素计算的；质量指标指数是以质量指标作为同度量因素计算的。（　　）

5．按照个体指数和报告期销售额计算的价格指数是加权算术平均数指数。（　　）

6．按说明对象的特征不同，指数可分为质量指标指数和数量指标指数。（　　）

7．价格上涨后，同样多的人民币报告期所购买商品的数量比基期少 5%，因此价格上涨了 5%。（　　）

8．在指数体系中，总量指数与各因素指数之间的数量关系是总量指数等于各因素指数之和。（　　）

三、计算题

1．某地区居民消费价格指数为 103%，消费量指数为 110%，求该地区居民消费额总指数。

2．某地区社会商品销售额报告期为 10 亿元，比基期增加 9 亿元，物价指数涨了 3%，试分析报告期比基期的商品销售量的变动情况。

3．某商店商品销售额增长 2.9%，价格下降 2%，问销售量指数为多少。

4．某商场商品销售量及价格资料如表 6-6 所示。

表 6-6　某商场商品销售量及价格情况表

商　品	单　位	销 售 量		价格（元）	
		基　期	报 告 期	基　期	报 告 期
甲	件	600	540	1 250	1 300
乙	套	800	910	20	25
丙	台	750	800	6	5.6

试计算该商场商品的：

（1）销售量总指数；

（2）销售价格总指数；

（3）销售额指数。



课外延伸阅读

中国物流业景气指数（LPI）

中国物流业景气指数（LPI）主要由业务总量、新订单、从业人员、库存周转次数、设备利用率、平均库存量、资金周转率、主营业务成本、主营业务利润、物流服务价格、固定资产投资完成额、业务活动预期 12 个分项指数和一个合成指数构成。其中合成指数由业务总量、新订单、从业人员、库存周转次数、设备利用率 5 项指数加权合成，称为中国物流业景气指数，英文缩写为 LPI。物流业景气指数反映物流业经济发展的总体变化情况，以 50%作为经济强弱的分界点，高于 50%反映物流业经济扩张；低于 50%则反映物流业经济收缩。

2013 年 3 月 5 日中国物流与采购联合会、中国物流信息中心正式对外发布：今后，中国物流业景气指数将于每月 5 日上午 9 时通过媒体对外发布，发布官方网站为中国物流与采购联合会网站和中国物流信息中心网站。

2014 年和 2015 年中国物流业景气指数如表 6-7 所示。

表 6-7　2014 年和 2015 年中国物流业景气指数

年　份	月　份	物流业景气指数（%）	年　份	月　份	物流业景气指数（%）
2014	1	51.5	2015	1	56.3
	2	51.9		2	54.9
	3	53		3	58
	4	57.7		4	58.6
	5	55.2		5	58
	6	56.7		6	58
	7	56.8		7	52.2
	8	54.1		8	52
	9	56.4		9	52.2
	10	54.9		10	53.5
	11	56.5		11	54.2
	12	57.5		12	55

物流业景气指数调查采用 PPS（Probability Proportional to Size）抽样方法，按照各物流行业对物流业主营业务收入的贡献度确定各行业的样本数。在此基础上兼顾样本的区域分布、企业类型分布、规模分布。物流业景气指数调查问卷涉及业务总量、新订单、库存周转次数、设备利用率、从业人员、平均库存量、资金周转率、主营业务成本、主营业务利润、物流服务价格、固定资产投资完成额、业务活动预期等 12 个问题。对每个问题分别计算扩散指数，即正向回答的企业个数百分比加上回答不变的百分比的一半。

中国物流业景气指数调查结果基本反映了我国物流业发展运行的总体情况，与货运量、快递业务量、港口货物吞吐量等物流相关指标，与工业生产、进出口贸易、固定资产投资、货币投放等相关经济指标具有较高的关联性。中国物流业景气指数调查丰富了物流统计指标体系，有效弥补了现行物流统计的不足，增加了观察、预测、分析我国物流业运行发展趋势的新视角，为进一步加强物流运行与国民经济的关联性研究奠定了基础，为指导企业生产经营与投资等活动提供了依据。

资料来源：http://www.clic.org.cn/wljqzs/index_3.jhtml.



第7章 抽样推断

学习目标要求

1. 理解抽样推断的基本概念、特点和作用。
2. 掌握样本容量的确定方法；利用样本资料来推断总体数量特征的原理及方法。
3. 了解抽样误差产生的原因；理解抽样平均误差、抽样极限误差、概率度和概率保证程度之间的关系。
4. 掌握点估计和区间估计的方法。

重点难点提示

1. 抽样误差的计算与实际应用。
2. 区间估计方法。

教学引导案例

国家统计局城市司高级统计师刘建伟解读2016年9月、10月房价数据

国家统计局今日发布了70个大中城市住宅销售价格统计数据。对此，国家统计局城市司高级统计师刘建伟进行了解读。

一、一线和部分热点二线城市有针对性地实施调控政策，10月份房地产市场明显降温

为快速反映近期房地产市场的变化，国家统计局在做好9月份70个大中城市房价统计的同时，对其中15个一线和部分热点二线城市10月上半月的房地产市场情况进行了统计。结果显示，一线和部分热点二线城市因地制宜、因城施策实施有针对性的调控政策后，房地产市场出现积极变化。从房价环比涨跌幅度看，15个一线和部分热点二线城市房价过快上涨的势头得到明显遏制，房价走势趋稳。与9月比，一线和部分热点二线城市10月上半月新建商品住宅价格环比指数均有所降低。其中，深圳和成都10月上半月新建商品住宅价格环比出现下降，降幅分别为0.3%和0.1%；其他城市10月上半月新建商品住宅价格环比涨幅均明显回落，回落幅度在1.0%～3.8%。从新建商品住宅网签成交量看，一线和部分热点二线城市中，多数城市10月上半月的成交量比9月下半月明显下降。其中，成交量环比下降80%～60%的城市有4个，下降60%～40%的城市有3个，下降40%～20%的城市有3个。

二、9 月份 70 个大中城市房价分化明显

9 月份，70 个大中城市中，一线和部分热点二线城市与其他城市的房价走势有明显差异。从环比看，一线和部分热点二线城市中，新建商品住宅价格涨幅在 2%以上的城市有 14 个；其余城市中，新建商品住宅价格环比下降的城市有 6 个，持平的城市有 1 个，涨幅在 2%及以内的城市有 40 个。从同比看，一线和部分热点二线城市中，新建商品住宅价格涨幅在 10%以上的城市有 14 个；其余城市中，新建商品住宅价格同比下降的城市有 6 个，涨幅在 10%以内的城市有 42 个。

资料来源：国家统计局

http://www.stats.gov.cn/tjsj/sjjd/201610/t20161021_1412286.html

思考：

1. 案例中提供了大量的统计数据，这些数据是对全国抽样 70 个大中城市进行调查得到的。这些统计数据是怎样计算出来的？

2. 抽样调查中，样本是如何抽取的？抽样误差是怎样控制的？

第一节　抽样推断概述

一、抽样推断

（一）抽样推断的概念与特点

1. 抽样推断的概念

抽样推断是按照随机原则从总体中抽取一部分总体单位作为样本单位，组成样本总体，并根据样本的数量特征对总体的数量特征做出具有一定可靠程度的估计和推断的统计分析方法。

随机原则又称同等可能性原则，即在总体中抽取样本时，每一个单位被抽中的机会都相等，被抽中与不被抽中纯属偶然，不受人的主观因素影响。按随机原则取样，能使抽取出来的样本分布近似于总体的分布，使样本的代表性增强。

抽样推断包含了两个含义：抽样调查和统计推断。

抽样调查是一种非全面调查，按随机原则从总体中抽出一部分单位进行调查以获得有关的数据资料，目的是推断总体。

统计推断是根据抽样调查所获得的样本信息对总体的数量特征做出具有一定可靠程度的估计和判断。

2. 抽样推断的特点

抽样推断具有以下特点。

（1）抽样推断是用样本指标值来估计总体指标值

由样本的已知资料去估计未知的总体特征，或者说，由部分去认识总体的全面情况，这

是抽样调查与其他非全面调查显著不同的方面之一。非全面调查中的重点调查通过对少数重点单位进行了解，只能实现对总体情况的基本认识，而不是全面的认识。

与全面调查相比，抽样调查只需要调查总体中的部分单位而非全部单位，因而是一种非全面调查。需要注意的是，抽样所抽取的全部样本单位作为一个“代表团”，用整个“代表团”来代表总体，而不是用随意挑选的个别单位代表总体。

（2）抽样的随机原则是抽样推断的前提

只有严格遵循随机原则才能使调查对象中某个单位都有同等机会被抽中或抽不中，以保证被抽中单位的次数分布类型与调查对象（总体）的分布状况接近，从而增在被抽中单位对总体的代表性。只有严格遵循随机原则，才能计算抽样误差，才能使用抽样误差去推断总体指标数值时估计出有多大把握。

但是，随机性不等于随意性，因为随意性受人的主观因素影响。

（3）抽样推断的误差是可以事先计算并加以控制的

利用概率论理论可以事先计算出抽样误差，并通过各种组织措施来控制抽样误差范围，保证抽样推断结果达到预期可靠程度的要求。

由于抽样调查具有随机性这一特点，从而使其可以建立在专门研究随机事件规律的概率论基础之上，运用概率论原理和推断的可靠程度，使抽样误差控制在一定的范围内。这一特点是其他非全面调查所不具备的。值得注意的是，抽样误差虽可以控制但不可以消除。

（二）抽样推断的作用

抽样推断作为一种科学、经济的统计方法，在科学研究、社会经济管理等方面已经得到普遍应用。从原则上讲，为取得大量社会经济现象数量方面的统计资料，在许多场合都可以运用抽样推断的方法，在某些特殊场合，甚至还必须运用抽样推断的方法。在统计实践中，抽样推断主要表现出以下几个方面的作用。

1. 适用于调查具有破坏性的场合

调查具有破坏性是指被抽取的单位经过观测之后就失去了原有的形态或功能。对这种情况的调查,只能采用抽样方法。例如，某灯泡厂在某时期生产了一大批灯泡，现在要了解这批灯泡的平均使用寿命。若采用全面调查，须将这批灯泡全部通电直至全部报废为止，那该灯泡厂的生产成果也就化为乌有了。像这种具有破坏性的调查，只能采用抽样调查。

2. 对无限总体或总体规模非常大的场合进行调查

例如,要测定湖水的含氧量；对大气污染程度的调查；对太空中行星运动规律的研究；对森林木材积蓄量的调查等，也只能采用抽样调查方法。

3. 不必要进行全面调查但又需要知道总体全面情况

例如，对居民家庭收支情况的调查；对成人身高和健康状况的调查；对人口流动情况的调查等。

4. 用于对全面调查的结果进行核查和修正

全面调查由于涉及面广，参加人员多，较容易产生漏查、重复、填错、计算错误等调查性差错，而抽样调查由于抽取的单位少，有准确度高及抽样误差可以控制等优点，因此，全面调查尤其是普查一旦结束就需要通过抽样来检验其质量，并进一步校正原来的数字。

5. 用于资料时效性要求很强的场合

例如，稻谷等农产品收割前一般都要进行产量预计，以便及时地组织仓储和运输设备；对工厂连续大量生产的产品质量进行检验和控制；新产品投产有利与否的统计检验等，都对调查结果有很强的及时性要求，一般均需采用抽样调查方法。

6. 可以节约人力、物力、财力和时间

抽样推断与全面调查相比较，调查单位比全面调查少得多，可以节约很多资源，又能较快地得到调查结果，这对许多工作都是有利的。特别是对于总体范围很大、单位很多的现象总体，抽样推断更显优越性。

总之，抽样调查是一种科学实用的调查方法，不仅广泛应用于自然科学领域，也越来越多地应用于社会经济领域。

二、抽样推断的理论基础

抽样推断是建立在概率论的大数定律和中心极限定理基础上的科学推断方法。

1. 大数定律

大数定律又称大数法则，是随机变量出现的基本规律。大致内容是在对某一现象观察过程中由大量互相独立的随机变量构成的总体，受偶然因素的影响，每次所得的结果不同，但经过大量观察并加以综合平均后，消除了偶然因素引起的差异，而接近于总体的平均值，使现象总体某一标志的规律及其共同特征在数量上、质量上显示出来。因此，大数定律为统计学研究现象在数量方面和数量关系上提供了科学依据。

2. 中心极限定理

中心极限定理是阐述平均数分布状态的理论，即不管总体的分布呈何种状态，由一系列随机样本组成的所有可能平均数的分布均服从正态分布。这为抽样推断提供了极为重要的数学依据和有效方法，利用这种关系，我们只要能够从样本总体中计算出算术平均数和标注差，就可以推断并估算出全及总体的相应指标。

大数定律帮助我们认识了样本平均数趋近总体平均数的趋势，这是推断的前提条件；中心极限定理则可帮助我们正确测算样本平均数与总体平均数之间的误差和样本平均数推断总体平均数的可靠程度，这是推断的主要依据。

三、抽样推断中的几个基本概念

（一）全及总体

全及总体是指所要了解认识的对象的全体，即统计总体，简称为总体。总体是具有某种共同性质或特征的许多单位的集合体。全及总体的单位数通常用 N 来表示。总体可从不同角度进行分类。例如，按其包含的总体单位数多少分为有限总体和无限总体；按照全及总体中各单位标志的性质不同，可以分为属性总体和变量总体。

（二）样本总体

样本总体是从全及总体中随机抽取出来，代表全及总体的那部分单位的集合体，又称抽样总体，简称样本或子样。样本总体的单位数称为样本容量，通常用 n 表示。一般来说，样本单位数达到或超过 30 个称为大样本，而在 30 个以下称为小样本。在对社会经济现象进行抽样调查时，多数采用大样本。

例如，某城市有 25 万个家庭，要采用抽样推断的方法研究其收支情况，则该市全部住户家庭构成全及总体，N=25 万。如果要从全部住户家庭中随机抽取 4‰，即 1000 户进行调查，则被抽中的 1000 户构成样本总体，则样本总体 n=1000。

对于一次抽样调查，全及总体是唯一确定的，而样本则是不确定的，一个全及总体可能抽出很多个样本总体，样本的个数和样本的容量有关，也和抽样的方法有关。

（三）全及指标

全及指标是根据全及总体各个单位的标志值或标志属性计算的，反映总体某种属性或特征的综合指标。常用的全及指标有全及平均数、全及成数、总体数量标志标准差及方差、总体是非标志标准差及方差。

1. 全及平均数

全及平均数又称总体平均数，是指全及总体各个单位的标志值的平均数，用 $\overline{X}$ 表示。

①在总体未分组的情况下：

$$\overline{X}=\frac{\sum_{i=1}^{N}X_i}{N}=\frac{\sum X}{N} \tag{7-1}$$

式中，X_i (i=1,2,3,⋯,N) 表示全及总体各单位标志值。

②在总体分组的情况下：

$$\overline{X}=\frac{\sum_{i=1}^{k}X_iF_i}{\sum_{i=1}^{k}F_i}=\frac{\sum XF}{\sum F} \tag{7-2}$$

式中，X_i (i=1,2,3,⋯,k) 表示总体各组的标志值；

F_i (i=1,2,3,⋯,k)表示总体各组分配次数；

k 表示将总体分组后分组的组数。

2. 全及成数

全及成数又称总体成数，它是指全及总体中具有某一相同标志表现的单位数占全及总体单位数的比重，用 P 表示。全及总体中具有某一相同标志表现的单位数用 n_1 表示，则 $P=n_1/N$。反之，如果将全及总体中不具有某一相同标志表现的单位数占全及总体单位数的比重用 Q 表示，全及总体中不具有某一相同标志表现的单位数用 n_0 表示，则 $Q=n_0/N$，且 P 与 Q 互为余数，即 $P+Q=1$ 或 $P=1-Q$ 或 $Q=1-P$。

3. 总体数量标志标准差及方差

①总体数量标志标准差是指全及总体中根据各单位标志值计算的标准差，记作 σ。

在总体未分组的情况下：

$$\sigma=\sqrt{\frac{\sum_{i=1}^{N}(X_i-\overline{X})^2}{N}} \tag{7-3}$$

在总体分组的情况下：

$$\sigma=\sqrt{\frac{\sum_{i=1}^{k}(X_i-\overline{X})^2F_i}{\sum_{i=1}^{k}F_i}} \tag{7-4}$$

②总体方差是总体标准差的平方，记作

$$\sigma^2=\frac{\sum_{i=1}^{N}(X_i-\overline{X})^2}{N} \tag{7-5}$$

或

$$\sigma^2=\frac{\sum_{i=1}^{k}(X_i-\overline{X})^2F_i}{\sum_{i=1}^{k}F_i} \tag{7-6}$$

（四）样本指标

样本指标又称抽样指标，它是根据样本总体各单位标志值或标志特征计算的，反映样本数量特征的综合指标。

样本指标与全及指标是相对应的，也可分为四种类型的六个指标，即样本平均数、样本成数、样本数量标志标准差及方差、样本是非标志标准差及方差。

1. 样本平均数

样本平均数又称抽样平均数，它是指样本总体各单位标志值的平均数，用 $\bar{x}$ 表示。

①在样本未分组的情况下：

$$\bar{x}=\frac{\sum_{i=1}^{n}x_i}{n} \tag{7-7}$$

式中，X_i (i=1,2,3,⋯,N) 表示样本各单位标志值。

②在样本分组的情况下：

$$\bar{x}=\frac{\sum_{i=1}^{k}x_i f_i}{\sum_{i=1}^{k}f_i} \tag{7-8}$$

式中，x_i (i=1,2,3,⋯,k) 表示样本各组的标志值；

f_i (i=1,2,3,⋯,k) 表示样本各组的分配次数；

k 表示将总体分组后，分组的组数。

2. 样本成数

样本成数又称抽样成数，它是指样本中具有某一相同标志表现的单位数占样本容量的比重，用 p 表示。将样本总体中不具有某一相同标志表现的单位数占样本容量的比重用 q 表示，则 p 与 q 互为余数，即 $p+q$=1 或 p=1−q 或 q=1−p。

3. 样本数量标志标准差及方差

样本数量标志标准差是指样本中根据各单位标志值计算的标准差，记作 S。样本方差是样本标准差的平方，记作 S^2。

①在样本未分组的情况下：

$$S=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}} \tag{7-9}$$

②在样本分组的情况下：

$$S=\sqrt{\frac{\sum_{i=1}^{k}(x_i-\bar{x})^2 f_i}{\sum_{i=1}^{k}f_i}} \tag{7-10}$$

（五）重复抽样

重复抽样又称重置抽样或放回抽样。它是指从全及总体 N 个单位中随机抽取容量为 n 的样本时，每次从总体中随机抽取一个样本单位，把结果登记下来后，重新将其放回到总体中去重新参加抽样，再从全及总体中抽取下一个样本单位，如此反复抽样、反复放回，直到抽完 n 个样本单位为止。

重复抽样方法具有如下特点：

其一，每次抽取都不改变总体的结构；

其二，每个单位在每次抽取中被抽中的概率均相等；

其三，总体中任一单位都有重复被抽中的可能。

（六）不重复抽样

不重复抽样也称不重置抽样、不回置抽样。它是指从总体 N 个单位中随机抽取容量为 n 的样本时，每次从总体中随机抽取一个样本单位，不再放回去，下一次则从剩下的总体单位中继续进行抽取，如此反复构成一个样本，就是说，每个总体单位只能被抽取一次，所以从总体中每抽取一次，总体就少一个单位，因此，先后抽出来的各个单位被抽中机会是不相等的。

不重复抽样方法具有如下特点：

其一，总体的内部结构会随每次抽取而改变；

其二，总体中任一单位最多只有一次被抽中的机会；

其三，每个单位在每次抽取中被抽中的概率是不相等的。

在抽样推断实践中，一般采用不重复抽样。主要有两个原因：其一，在某些抽样调查中，样本单位不可放回。其二，相对于重复抽样，不重复抽样条件下，样本的代表性较高，抽样误差较小。

第二节　抽 样 误 差

一、抽样误差

（一）抽样误差的概念

抽样误差是指按随机原则抽样时，由于调查范围的非全面性和抽样的随机性而产生的样本指标与总体指标之间的误差。它是随机抽样调查所特有的一种偶然性的代表性误差。在随机抽样调查中，抽样误差是不可克服、不可避免的，但可以通过一定的方法加以估计和控制。

（二）抽样误差的性质

1. 抽样误差属于随机误差

在抽样调查过程中，主要产生两种统计误差：一种是登记性误差，另一种是代表性误差。登记性误差是指在统计调查过程中，由于主客观原因在登记、记录、汇总、计算中产生的差错。只要有较强的工作责任心和较高的技术水平，登记性误差是完全可以避免的。

代表性误差产生的情况则不同。代表性误差又可分为系统性误差和抽样误差。系统性误差是由于破坏了或没有完全遵守随机原则所产生的代表性误差，而抽样误差是即使完全遵守了随机原则，仍然存在的代表性误差。由于在抽样调查理论中总是假定完全遵守随机原则，因而系统性误差为零。因此，抽样误差即指抽样调查的代表性误差，其可控性强，可以将其控制在一定的范围内，但不可消灭。

2. 抽样误差是抽样平均误差

抽样误差是指抽样平均误差，而不是抽样实际误差。由于在现实的抽样推断中无法找到一一对应的抽样实际误差，即每个样本指标与总体被估计的真实指标之间的离差。为了反映抽样误差的一般水平，需要采用抽样平均误差指标。抽样平均误差是指所有可能出现的样本指标的标准差。

抽样平均误差是指从总体中任意抽取一个样本所得的样本指标与总体指标的平均偏离程度。通俗地说，抽样平均误差的意思是：无论抽取哪个样本，平均说来会有多大的误差。

（三）影响抽样误差的因素

影响抽样误差大小的因素主要有以下几个方面。

1. 总体各单位标志值的变异程度

总体各单位标志值的变异程度即受全及总体的方差或标准差数值大小的影响。如果总体变异性小，则所抽取各样本的指标与总体指标的离差也较小，再求其平均而得到抽样误差也较小。因而，抽样误差的大小与总体的方差（或标准差）大小成正比例关系。

2. 样本单位数目的多少

抽取的样本单位数越多，那么它对总体的代表性越好，则抽样误差越小；反之，则越大。可见，抽样误差的大小与抽取的样本容量成反比。

3. 抽样方法

在其他条件相同的情况下，不重复抽样的抽样误差小于重复抽样的抽样误差。这是因为不重复抽样下对已抽过的总体单位不再放回参加下一次抽取，从而避免了重复中选，因此，不重复抽样的样本比重复抽样的样本更能反映总体的特征，故抽样误差较小。

4. 抽样调查的组织形式

抽样调查按其组织形式不同可分为简单随机抽样、分类抽样、等距抽样、整群抽样。相同条件下，不同的抽样组织形式有不同的抽样误差，这是因为按不同组织形式所抽取的样本对于总体的代表性不同。

二、抽样平均误差

（一）抽样平均误差的概念

抽样平均误差是指根据随机原则抽样时，所有可能出现的样本平均数（或成数）的标准差。抽样平均误差反映了样本指标与全及指标的平均离差程度，抽样平均误差的作用首先表现在它能够说明样本指标代表性的大小，抽样平均误差大，说明样本指标对全及指标的代表性低；反之则高。

抽样平均误差的理论公式为：

$$\mu_{\bar{x}}=\sqrt{\frac{\sum(\overline{x_i}-\overline{X})^2}{\text{所有可能样本数目}}} \tag{7-11}$$

式中，$\mu_{\bar{x}}$ 代表样本平均数的抽样平均误差；$\overline{x_i}$ 代表所有可能样本的平均数；$\overline{X}$ 代表总体平均数。

实践中，总体平均数是要推算的未知数，所有可能样本的平均数也是无法取得的。因此不能直接利用定义性公式计算抽样平均误差。

（二）抽样平均误差的实际计算方法

抽样平均误差的实际计算公式是由数理统计的有关原理和定理推导得出的。按照抽取样本单位的方式和方法的不同，抽样平均误差的实际计算方法也有所不同。抽样平均误差有平均数的抽样平均误差和成数的抽样平均误差两种，且都分别采用重复抽样和不重复抽样两种方法，因此，实际具体计算方法有所差别。

1. 重复抽样的抽样平均误差计算公式

①平均数的抽样平均误差：

$$\mu_{\bar{x}}=\sqrt{\frac{\sigma^2}{n}}=\frac{\sigma}{\sqrt{n}} \tag{7-12}$$

②成数的抽样平均误差：

$$\mu_p=\sqrt{\frac{p(1-p)}{n}} \tag{7-13}$$

2. 不重复抽样的抽样平均误差计算公式

①平均数的抽样平均误差：

$$\mu_{\bar{x}}=\sqrt{\frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)} \tag{7-14}$$

N 很大时，公式可近似地表示为：

$$\mu_{\bar{x}}=\sqrt{\frac{\sigma^2}{n}\left(1-\frac{n}{N}\right)} \tag{7-15}$$

②成数的抽样平均误差：

$$\mu_p=\sqrt{\frac{p(1-p)}{n}\left(\frac{N-n}{N-1}\right)} \tag{7-16}$$

N 很大时，公式可近似地表示为：

$$\mu_p=\sqrt{\frac{p(1-p)}{n}\left(1-\frac{n}{N}\right)} \tag{7-17}$$

从上述公式中可以看出，在其他条件不变的情况下，不重复抽样的抽样平均误差小于重复抽样的抽样平均误差，不重复抽样的样本代表性较大。当总体单位数 N 很大时，$\frac{n}{N}$ 很小，而 $1-\frac{n}{N}$ 就接近于 1，在这种情况下，不重复抽样的抽样平均误差近似于重复抽样的抽样平均

误差。

在计算抽样误差时，由于总体标准差是未知的，在大样本情况下，可用样本标准差代替总体标准差，也可以用历史标准差或试验标准差代替。

（三）抽样平均误差的实际应用

1. 平均数抽样平均误差的计算



【例 7-1】

从某厂生产的 20 000 只日光灯管中随机抽取 1 000 只进行检查，假如该产品平均使用寿命的标准差 S 为 10 小时，试根据重复抽样和不重复抽样的方法，计算该厂日光灯管平均使用寿命的平均误差。

（1）按重复抽样计算抽样平均误差

根据式（7-12）得

$$\mu_{\bar{x}}=\sqrt{\frac{S^2}{n}}=\sqrt{\frac{10^2}{1000}}=0.32\text{（小时）}$$

（2）按不重复抽样计算抽样平均误差

根据式（7-15）得

$$\mu_{\bar{x}}=\sqrt{\frac{S^2}{n}\left(1-\frac{n}{N}\right)}=\sqrt{\frac{10^2}{1000}\left(1-\frac{1000}{20\,000}\right)}=0.31\text{（小时）}$$



【例 7-2】

某工业企业有 5000 名工人，从其中随机抽取 100 人进行日常量调查，资料如表 7-1 所示。试计算平均日产量的抽样平均误差。

表 7-1　抽样平均误差计算表

按工人日产量分组 x（件）	工人数 f（人）	xf	$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
200	1	200	−5	25	25
201	1	201	−4	16	16
202	3	606	−3	9	27
203	10	2030	−2	4	40
204	20	1080	−1	1	20
205	30	6150	0	0	0
206	20	4120	1	1	20
207	10	2070	2	4	40
208	3	624	3	9	27
209	2	418	4	16	32
合计	100	20 499	—	—	247

首先计算样本平均数 $\bar{x}$ 和样本方差 S^2:

$$\bar{x}=\frac{\sum xf}{\sum f}=\frac{20\,499}{100}=205(\text{件})$$

$$S^2=\frac{\sum(x-\bar{x})^2 f}{\sum f}=\frac{247}{100}=2.47$$

然后计算抽样平均误差。

（1）按重复抽样计算抽样平均误差。

根据式（7-12）得

$$\mu_{\bar{x}}=\sqrt{\frac{S^2}{n}}=\sqrt{\frac{2.47}{100}}=0.157\text{（件）}$$

（2）按不重复抽样计算抽样平均误差。

根据式（7-15）得

$$\mu_{\bar{x}}=\sqrt{\frac{S^2}{n}(1-\frac{n}{N})}=\sqrt{\frac{2.47}{100}(1-\frac{100}{5000})}=0.156\text{（件）}$$

2. 成数抽样平均误差的计算



【例 7-3】

从某厂生产的 10 000 件产品中随机抽取 100 件进行调查，测得有 1 件不合格。试求产品合格率的抽样平均误差。

根据条件可知，合格率 $P=\frac{100-1}{100}\times 100\%=0.99$。

（1）在重复抽样条件下:

$$\mu_p=\sqrt{\frac{p(1-p)}{n}}=\sqrt{\frac{0.99\times 0.01}{100}}=1\%$$

（2）在不重复抽样条件下:

$$\mu_p=\sqrt{\frac{p(1-p)}{n}(1-\frac{n}{N})}=\sqrt{\frac{0.99\times 0.01}{100}(1-\frac{1}{100})}=1\%$$

计算结果表明，当 $\frac{n}{N}$ 较小时，重复抽样和不重复抽样的抽样平均误差几乎没有差别。

三、抽样极限误差

抽样极限误差是指样本指标和全及指标之间抽样误差的最大可能范围。由于全及指标是一个确定的数，而样本指标则是围绕着全及指标左右变动的量，它与全及指标可能产生正离差，也可能产生负离差，样本指标变动的上限或下限与全及指标之差的绝对值就可以表示抽样误差的可能范围，这种以绝对值形式表示的抽样误差可能范围称为抽样极限误差。抽样极

限误差用符号Δ表示。

样本平均数的极限误差：

$$\Delta_{\bar{x}} = \left|\bar{x} - \overline{X}\right| \tag{7-18}$$

样本成数的极限误差：

$$\Delta_p = \left|p - P\right| \tag{7-19}$$

上述公式用不等式表示：

$$\bar{x} - \Delta_{\bar{x}} \leqslant \overline{X} \leqslant \bar{x} + \Delta_{\bar{x}}$$

$$p - \Delta_p \leqslant P \leqslant p + \Delta_p$$

从不等式中可看出，只要知道样本平均数 $\bar{x}$ 和样本成数 p 及抽样极限误差 $\Delta_{\bar{x}}$ 和 Δ_p，就可以估计全及总体平均数 $\overline{X}$ 和全及总体成数 P 所在的可能范围。例如某地区 1000 户居民人均年收入为 2.5 万元，不重复抽样的抽样平均误差为 0.026 万元，如果确定极限误差为 0.022 万元，则该地区 1000 户居民的人均年收入的可能范围是在（2.5±0.022）万元，即 24 780～25 220 元。

第三节　总体指标的估计方法

抽样推断就是根据所计算的样本指标来推断相应的全及指标。由于这种推断不完全准确，实际上是一种科学的估计，因此抽样推断方法又称为抽样估计方法，包括点估计和区间估计两种方法。由于全及指标是表明总体数量特征的参数，所以这种估计也可以称为参数估计。

一、点估计

点估计也称定值估计，它是根据样本资料得出的样本指标，直接用以代表相应全及指标，即 $\bar{x} = \overline{X}$，$p = P$。

点估计是用某一样本指标直接推断相应的总体指标的估计形式。例如，根据样本的平均数估计总体的平均数、根据样本的成数估计总体的成数等。

例如，对 10 000 件产品按 3%的比例进行随机抽样检验，测得 300 件产品的平均使用寿命是 5000 小时，合格率为 96%，则可以认为 10 000 件产品的总平均使用寿命是 5000 小时，合格率为 96%。

点估计的优点是简便易行、原理直观；不足之处是没有考虑到抽样误差，更没有指出误差在一定范围内的概率保证程度有多大。点估计没有考虑估计的精度，也没有考虑估计的可靠程度，正是这种局限性决定了在实际工作中对总体指标的估计习惯于用区间估计。

二、区间估计

（一）区间估计的概念

区间估计是根据样本指标和抽样误差结合推断的可靠程度来估计总体的可能区间的方法。进行区间估计时必须掌握以下基本要点。

1. 根据样本指标和抽样极限误差估计全及指标的可能范围

即 $\bar{x}-\Delta_{\bar{x}} \leqslant \overline{X} \leqslant \bar{x}+\Delta_{\bar{x}}$ 或 $\bar{x}-\Delta_p \leqslant \overline{X} \leqslant \bar{x}+\Delta_p$。由抽样极限误差估计全及指标的范围称作置信区间。总体平均数的置信区间记作（$\bar{x}-\Delta_{\bar{x}}, \bar{x}+\Delta_{\bar{x}}$）；总体成数的置信区间记作（$p-\Delta_p, p+\Delta_p$）。

2. 区间估计所所表明的只是一个可能范围，不是一个绝对可靠范围

由于全及指标是固定的、未知的，而用样本构造的区间则是不固定的。我们抽取不同的样本，用该方法可以得到不同的区间，从这个意义上说，置信区间是一个随机区间，它会因样本的不同而不同，而且不是所有的区间都包含全及指标的真值。

3. 要提高估计的可靠程度，必须扩大误差范围

抽样误差范围决定估计的准确性，而概率保证度决定估计的可靠性。

利用一个样本对总体进行估计时，若要提高估计的准确性，必然会降低估计的可靠性。因此在抽样估计时，只能对其中一个要素提出要求，来推算另一个要素的变动情况。例如，对估计的准确性提出要求，即要求误差范围不超过给定的标准，来推算估计的可靠性。或者对估计的可靠性提出要求，即要求给出一定的概率保证度来推算可能的误差范围。根据给定的条件不同，有两种估计方法。

（二）区间估计的方法

1. 根据给定的抽样误差范围求概率保证程度

具体步骤如下。

第一步，抽取样本，计算样本平均数或样本成数，作为全及指标的估计值，并计算样本标准差以推算抽样平均误差。

第二步，根据给定的抽样极限误差范围估计全及指标的上限和下限。

第三步，将抽样极限误差除以抽样平均误差，求出概率度 t，再根据 t 值查正态分布概率表求出相应的概率保证程度 $F(t)$，进而对全及指标做出区间估计。



【例 7-4】

对一批某型号的电子元件进行耐用性能检查，按重复随机抽样的资料分组，如表 7-2 所示，要求耐用时数的允许误差范围不超过 10.5 小时，试估计该批电子元件的平均耐用时数。

表 7-2　某型号的电子元件抽查资料表

耐用时数（小时）	组中值（x）	元件数（f）	xf
900 以下	875	1	875
900～950	925	2	1850
950～1000	975	6	5850
1000～1050	1025	35	35 875
1050～1100	1075	43	46 225
1100～1150	1125	9	10 125
1150～1200	1175	3	3525
1200 以上	1225	1	1225
合计	—	100	105 550

①计算样本平均数、样本标准差、抽样平均误差：

$$\bar{x}=\frac{\sum xf}{\sum f}=\frac{105\,550}{100}=1055.5$$

$$s=\sqrt{\frac{\sum(x-\bar{x})^2 f}{\sum f}}=51.91\text{（小时）}$$

$$\mu_{\bar{x}}=\frac{s}{\sqrt{n}}=\frac{51.91}{\sqrt{100}}=5.191\text{（小时）}$$

②根据给定的允许误差范围 $\Delta_{\bar{x}}$=10.5 小时计算总体平均数的上下限：

$$\text{下限}=\bar{x}-\Delta_{\bar{x}}=1055.5-10.5=1045\text{（小时）}$$

$$\text{上限}=\bar{x}+\Delta_{\bar{x}}=1055.5+10.5=1066\text{（小时）}$$

③求概率保证程度 $F(t)$：

$$t=\frac{\Delta_{\bar{x}}}{\mu_{\bar{x}}}=\frac{10.5}{5.191}=2.02$$

根据 t=2.02，查正态分布概率表得概率保证度 $F(t)$=0.9566。

推断的结论是：在 95.66%的概率保证程度下，估计该批电子元件的耐用时数在 1045～1066 小时。



【例 7-5】

承例 7-4，假设该厂的产品质量检验标准规定，元件耐用时数达 1 000 小时以上为合格品，要求合格率的可允许误差范围不超过 4%，试估计该批电子元件的合格率。

①计算样本合格率、标准差、合格率的抽样平均误差：

根据表 7-2 可知，耐用时间达到 1000 小时以上的元件有 91 件，则

$$p=\frac{n_1}{n}=\frac{91}{100}=91\%$$

$$s_p^2=p(1-p)=0.91\times 0.09=0.0819$$

$$\mu_p=\sqrt{\frac{p(1-p)}{n}}=\sqrt{\frac{0.0819}{100}}=2.86\%$$

②根据该给定的允许误差范围 Δ_p=4%，求总体合格率的上下限：

$$\text{下限}=p-\Delta_p=91\%-4\%=87\%$$

$$\text{上限}=p+\Delta_p=91\%+4\%=95\%$$

③求概率保证程度 $F(t)$

$$t=\frac{\Delta_p}{\mu_p}=\frac{4\%}{2.86\%}=1.4$$

根据 t=1.4，查正态分布概率表得概率 $F(t)$=83.85%。

结论是：在 83.85%的概率保证程度下，估计该批电子元件的合格率在 87%～95%。

2. 根据给定概率保证程度并求抽样误差范围

具体步骤如下。

第一步，抽取样本，计算样本平均数或样本成数，作为全及指标的估计值，并计算样本标准差以推算抽样平均误差。

第二步，根据给定的概率保证度 $F(t)$ 的要求，查正态分布概率表求得概率度 t 值。

第三步，根据概率度 t 和抽样平均误差推算抽样极限误差，并根据抽样极限误差求出被估计总体的上下限。



【例 7-6】

对我国某城市进行居民家庭人均旅游消费支出调查，随机抽取 400 户居民家庭，调查得知居民家庭人均年旅游消费支出为 930 元，标准差为 200 元，要求以 89.9%的概率保证程度，估计该市人均年旅游消费支出额。

①根据抽样资料可知：n=400 户，$\bar{x}$=930 元，S=200 元，因此：

$$\mu_{\bar{x}}=\frac{s}{\sqrt{n}}=\frac{200}{\sqrt{400}}=10\text{（元）}$$

②根据 $F(t)$=89.9%，查正态分布概率表得 t=1.64。

③求抽样极限误差及被估计总体的上下限：

$$\Delta_{\bar{x}}=t\mu_{\bar{x}}=1.64\times10=16.4$$

$$\text{下限}=\bar{x}-\Delta_{\bar{x}}=930-16.4=913.6\text{（元）}$$

$$\text{上限}=\bar{x}+\Delta_{\bar{x}}=930+16.4=946.6\text{（元）}$$

结论：在 89.9%的概率保证程度下，估计该市居民家庭年人均旅游消费支出额在 913.6～946.4 元。



【例 7-7】

某市电视台为了解观众对某电视栏目的喜爱程度，在该市随机对 900 名居民进行调查，结果有 540 名观众喜欢该电视栏目，要求以 95%的概率保证程度估计该市居民喜欢该电视栏

目的比率。

①根据抽样资料可知:

$$p=\frac{n_1}{n}=\frac{540}{900}=60\%$$

$$s_p^2=p(1-p)=0.6\times0.4=0.24$$

$$\mu_p=\sqrt{\frac{p(1-p)}{n}}=\sqrt{\frac{0.24}{900}}=1.63\%$$

②根据 $F(t)$=95%，查正态分布概率表得 t=1.96。

③求抽样极限误差及被估计总体的上下限:

$\Delta_p=t\mu_p$ =1.96 × 1.63%=3.19%

下限 = $p-\Delta_p$ = 60%−3.19% = 56.81%

上限 = $p+\Delta_p$ = 60% + 3.19% = 63.19%

结论：在 95%的概率保证程度下，估计该市居民对此电视栏目喜爱的比率在 56.81%～63.19%。

第四节　样本容量确定与抽样方案设计

一、样本容量的确定

必要样本容量是指为了完成抽样调查任务，满足抽样调查的各项要求而科学计算的需要抽取的样本单位数。确定必要的样本容量是抽样调查方案中的一个重要问题。样本容量过大会增加调查费用，花费更多的人力和时间，不能充分发挥抽样调查的优越性。样本容量过小，样本代表性不足，抽样误差增大，对全及指标的推断必然不准确，失去实际价值。为了避免样本容量的过大或过小，必须恰当确定样本容量。

（一）影响必要样本容量的因素

1. 全及总体各单位标志变异程度的大小

全及总体各单位标志值之间的差异越大，方差也越大，因而，要求抽取的样本单位数就越多；反之，要求抽取的样本单位数就越少。例如，要了解某城市人均收入状况，如果人均收入水平相差较大，必要抽取的样本单位就要多些。

2. 允许的抽样极限误差的大小

对抽样推断的准确性要求越高，其允许的抽样极限误差越小，因而，要求抽取的样本单位数就越多；反之，允许的抽样极限误差越大，必要抽取的样本单位就越少。当允许误差为零时，抽样调查就变成了全面调查。

3. 抽样推断的可靠程度

如果要求抽样推断的可靠程度较高，即概率保证度 $F(t)$较大，则概率度 t 值也大，因而，要求抽取的样本单位数就较多；反之，要求的可靠程度越低，t 值也越小，则应抽取的样本单位数越少。

4. 抽样方法和抽样组织方式

不同的抽样方法和抽样组织方式会产生不同的抽样平均误差，进而影响推断估计的精度和可靠程度。例如，在其他条件相同的情况下，重复抽样比不重复抽样要抽取多一些样本单位；采用分类抽样的样本容量要小于简单随机抽样的样本容量。

（二）简单随机抽样必要样本容量的计算公式

1. 重复抽样条件下

①样本平均数的必要样本容量为：

$$n_{\bar{x}}=\frac{t^2\sigma^2}{\Delta^2_{\bar{x}}} \tag{7-20}$$

②样本成数的必要样本容量为：

$$n_p=\frac{t^2p(1-p)}{\Delta_p^2} \tag{7-21}$$

2. 不重复抽样条件下

（1）样本平均数的必要样本容量为：

$$n_{\bar{x}}=\frac{t^2\sigma^2N}{\Delta^2_{\bar{x}}N+t^2\sigma^2} \tag{7-22}$$

（2）样本成数的必要样本容量为：

$$n_p=\frac{t^2p(1-p)N}{\Delta_p^2N+t^2p(1-p)} \tag{7-23}$$

（三）简单随机抽样样本容量确定实例

【例 7-8】

对某油田的 2000 口油井的年产油量进行抽样调查。根据历史资料可知，油井年产油量的标准差为 200 吨，若要求抽样误差不超过 15 吨，概率保证程度为 95.45%，问应抽取多少口油井进行调查。

根据概率保证度 $F(t)$= 95.45%，查正态分布概率表得 t=2。

①重复抽样条件下:

$$n_{\bar{x}}=\frac{t^2\sigma^2}{\Delta^2_{\bar{x}}}=\frac{2^2\times200^2}{15^2}=712\text{（口）}$$

即在重复抽样条件下需抽查 712 口油井。

②不重复抽样条件下:

$$n_{\bar{x}}=\frac{t^2\sigma^2N}{\Delta^2_{\bar{x}}\mathrm{N}+t^2\sigma^2}=\frac{2^2\times200^2\times2000}{15^2\times2000+2^2\times200^2}=525\text{（口）}$$

即在不重复抽样条件下需抽查 712 口油井。



【例 7-9】

对某工厂的一批机械零件的合格率进行抽样调查，根据过去的资料，合格率曾有过 94%、97%和 99%三种情况，现要求允许误差不超过 1%，要求推断的把握程度为 95.45%，问至少要抽取多少个零件进行调查。

根据概率保证度 $F(t)$= 95.45%，查正态分布概率表得 t=2，Δ_p=0.01，由三个值分别计算的方差是：0.94 × 0.06=0.0564；0.97 × 0.03=0.0291；0.99 × 0.01=0.0099，因此应选最大的值为方差，即选 p=0.94。

故

$$n_p=\frac{t^2p(1-p)}{\Delta_p^2}=\frac{2^2\times0.94\times0.06}{0.01^2}=2256\text{（个）}$$

即应至少抽取 2256 个零件进行调查。



【例 7-10】

某电池厂对某种型号的电池进行电流强度检验。根据以往正常生产的经验，电池电流强度的标准差为 0.4 安培，合格率为 90%。现以重复抽样方式计算抽样数目，并要求在 95.45%的概率保证下，平均数的允许误差不得超过 0.08 安培，合格率的允许误差不得超过 5%，问应抽取多少只电池进行检验。

根据概率保证度 $F(t)$= 95.45%，得 t=2；σ=0.4，$\Delta_{\bar{x}}$=0.08，p=0.9，Δ_p=0.05。

根据以上资料，分别计算平均数样本数和成数样本数:

$$n_{\bar{x}}=\frac{t^2\sigma^2}{\Delta^2_{\bar{x}}}=\frac{2^2\times0.4^2}{0.08^2}=100\text{（只）}$$

$$n_p=\frac{t^2p(1-p)}{\Delta_p^2}=\frac{2^2\times0.9\times0.1}{0.05^2}=144\text{（只）}$$

根据计算结果，两个样本指标所必需的抽样数目不同，故这时应选用稍大的样本容量即 144 只，以满足二者的共同需要。

（四）确定必要样本容量应注意的问题

①上述计算样本容量公式所计算出的样本单位数是为达到所给定要求而必须抽取的最低样本单位数目。实际调查时可对计算的必要抽样数目进行调整。②当总体单位数不大时，如果采用不重复抽样的方法抽取样本，必须应用不重复抽样的计算公式计算必要抽样数目；当总体单位数很大时，虽然采用不重复抽样方法，但仍可采用重复抽样的计算公式计算必要抽样数目。③一个总体往往同时计算抽样平均数和抽样成数。由于它们的方差和允许误差范

围不同，因此，需要的必要抽样数目也不相同。为了防止由于样本单位数不足而扩大抽样误差，在实际工作中往往根据比较大的必要抽样数目进行抽样，以满足共同的需要。④如果有几个方差可以选用，宜选择最大数值的方差。⑤计算的样本容量如果带有小数，一般不采取四舍五入的方法取整数，而是用比这个数大的邻近整数代替。

二、抽样方案设计

抽样推断是以有效取得各项实际资料为基础的，要保证抽样估计的准确性和可靠性，必须结合一定的抽样调查组织方式做好抽样设计工作。根据随机抽样的原则，结合具体研究对象的性质及调查工作的目的和条件，在实践中，主要采用简单随机抽样、分类抽样、等距抽样和整群抽样四种抽样调查组织方式。

（一）简单随机抽样

1. 简单随机抽样的概念与特点

简单随机抽样也叫纯随机抽样，是指在进行抽样时，对全及总体不经过任何形式的整理和加工，直接从总体中随机抽取样本单位的抽样组织方式。简单随机抽样又分为重复抽样和不重复抽样两种方法。简单随机抽样是最简单、最基本的抽样组织方式。

从理论上说，这种抽样方式也是最符合随机原则的。这种方法的优点是使用起来简便易行，它适用于总体单位不多、各单位均匀分布于总体的各个部分且总体的各个部分都是同等分布的情况。

2. 简单随机抽样的方法

在进行简单随机抽样前应先确定总体范围，并对总体单位进行编号，形成明确的“抽样框”，即总体单位的名单，然后用抽签、摇号的方法或根据随机数表来抽选必要的样本单位数。

随机数表是指含有一系列随机数字的表格。这种表格的编制即可以借助电子计算机产生，也可以采用数码机产生或自己编制。表中数字的出现及其排列是随机形成的。使用此表时，可以竖查、横查、顺查、逆查，可以用表中每组数字左边的前几个数，也可以用中间或右边的某几位数，总之，可以随意使用。但一经确定使用某种方法后，中途不能改用其他方法，以保持在使用方法上的一致性。

查到在抽样编码范围内的号码就是被抽中的样本单位，直到抽取出足够的样本单位数为止。如果采用不重复抽样，应将所碰到的重复号码跳过去，按上面的方法继续抽选，直至样本单位抽满为止。

（二）分类抽样

1. 分类抽样的概念与特点

分类抽样又称类型抽样或分层抽样，它是先将全及总体按某个主要标志划分为几个类型组，然后在各组中按随机原则抽取样本单位的一种抽样组织方式。分类抽样的特点是把统计分组方法同贯彻随机原则结合起来。

分类抽样通过分组使同组内各单位之间的差异变小，并且保证每组有同等被抽选的机会，使样本结构趋近于总体结构。因此，分类抽样使从各组抽取的样本代表性较高，同时也降低了影响抽样平均误差的方差。

2. 分类抽样的方法

经过划类分组后，确定各类型组抽样单位数一般有两种方法。

第一种方法是各组所抽选的单位数按各组标志值的变动程度来确定，变动程度大的组，抽样数目可适当多一些；变动程度小的组，抽样数目可适当少一些，各组抽样数目没有统一的比例关系，这就是不等比例抽样。

第二种方法是按照等比例抽样分配各组的抽样单位数，即按各组单位数占总体单位总数的比重来分配抽样数目。例如，某市有 10 万户居民，按 1‰的比例抽取样本户进行居民家庭收入水平的抽样调查，则抽取样本户数为 100 户，若收入水平高的居民户占 20%，收入水平居中的居民户占 70%，收入水平低的居民户占 10%，则三个组所分配的样本户分别为：20 户、70 户和 10 户。

等比例抽样是在总体单位标志值均匀分布在全及总体各单位之间时使用的方法，而不等比例抽样则是在总体单位标志值分布不均衡、相差很大的情况时使用的方法。不等比例抽样多指某类型组的单位数在总体单位总数中占的比重过小时，如果按等比例抽样有可能一个样本单位得难抽到或即使能抽到样本单位，但份额太小，为了保证样本中各个类型组的单位数的代表性，只能采用不等比例抽样方式。

（三）等距抽样

1. 等距抽样的概念

等距抽样又称机械抽样或系统抽样，它是先将总体单位按某一标志顺序排列，然后按照固定顺序和相同间隔来抽取样本单位的抽样组织方式。作为总体各单位顺序排列的标志，可以是无关标志，也可以是有关标志。

无关标志是指排列标志与单位标志值的大小无关或不起主要影响作用，如时间、地理位置、门牌号码、姓氏笔画标志等。有关标志是指作为排队顺序的标志与单位标志值的大小有密切关系。

2. 等距抽样的方法

按样本单位抽选的方法不同，等距抽样分为随机起点等距抽样、半距起点等距抽样和对称等距抽样三种。

（1）随机起点等距抽样

先将总体单位按某种顺序排队，随机确定一个起点抽取第一个样本单位，然后每隔一定间隔抽取一个样本单位，直至样本单位抽满为止。

（2）半距起点等距抽样

这种方法要求各个样本单位都选在各组的中点位置上，其优点在于简单易行，而且当总体按有关标志排队时，能保证样本有充分的代表性。但半距起点等距抽样也有一定的局限性，表现在两个方面：首先，随机性不明显，若总体排队顺序确定了，样本容量也确定了，样本

单位也随之确定；其次，由于所抽样本的位置已被等距取样所固定，所有样本单位也被等距取样所固定，不能更改，不能进行样本轮换，抽样框的利用率过低。

（3）对称等距抽样

对一个调查总体，当样本单位抽取间隔 d 确定后，按这种方法要求，首先在第一组随机抽取第一个样本单位，假设该样本单位在全及总体单位中排的序号为 a；接着在第二组与第一个样本单位对称的位置上抽取第二个样本单位，其序号为 $2d-a$；又在第三组与第二个样本单位对称的位置上抽取第三个样本单位，其序号为 $2d+a$；以后陆续抽出的样本单位序号依次为 $4d-a$，$4d+a$，$6d-a$，$6d+a$，$8d-a$，$8d+a$…

对称等距抽样不仅简便易行，能够保证样本有充分的代表性，且随机性扩大，并能进行样本轮换，提高了抽样框的利用率。

（四）整群抽样

1. 整群抽样的概念

整群抽样与前几种抽样组织方式的最大区别在于，它的抽样单位不是单个的个体，而是成群的个体。整群抽样是将总体各单位划分成若干群，然后以群为单位，从中随机抽取一些群；对中选群的所有单位进行全面调查的抽样组织方式。

例如，按产品产出时间顺序，每隔一个固定的时间长度，抽一段时间的全部产量，即为一群；又如，对某市居民生活水平的调查可从该市全部居民委员会中随机抽选若干居委会，然后对被抽中的居委会的所有居民户进行深入的全面调查，再根据其调查结果推断该市全部居民生活水平状况，即是整群抽样。

2. 整群抽样的特点

整群抽样与分类抽样对比，虽然两者都将总体划分许多组，但划分组的作用不同。分类抽样划分的组称为“类”，它的作用是缩小总体，使总体的差异减少，抽取的样本仍是总体单位；整群抽样划分的组是“群”，它的作用是扩大单位。

整群抽样的优点是调查工作的组织和进行比较方便，缺点是调查的总体单位集中在少数样本群中，如果调查单位在总体中的分布不均匀，则抽样误差较大，代表性较低。因此，在群间差异性不大或不适宜单个抽选样本单位的情况下，可采用这种方式。



本章小结

抽样推断是按照随机原则从总体中抽取一部分总体单位作为样本单位，组成样本总体，并以样本的数量特征对总体的数量特征做出具有一定可靠程度的估计和推断的统计分析方法。抽样推断中常用的概念有：全及总体和样本总体、全及指标和样本指标、抽样平均误差和抽样极限误差、重复抽样和不重复抽样等。

抽样推断方法包括点估计和区间估计两种方法。在统计实践中，抽样的组织方式主要有：

简单随机抽样、分类抽样、等距抽样和整群抽样四种。影响必要样本容量的因素有：总体各单位标志变异程度、允许误差、概率度、抽样方法和抽样组织方式等。



复习思考题

1．什么是随机原则？在抽样调查中为什么要坚持随机原则？

2．抽样推断有哪些特点和作用？

3．什么是全及总体？什么是样本总体？各有哪些指标？

4．抽样平均误差和抽样极限误差有何关系？

5．点估计和区间估计两种抽样推断方法有何异同？

6．什么是必要样本容量？影响必要样本容量的因素有哪些？



实训基础训练

一、单项选择题

1．抽样推断是按照（　　）从总体中抽取一部分总体单位作为样本单位的。

A．准确性原则　　B．随机原则　　C．代表性原则　　D．可靠性原则

2．反映样本指标与总体指标之间的抽样误差可能范围的是（　　）。

A．抽样极限误差　　B．概率度　　C．概率保证程度　　D．抽样平均误差

3．在实际工作中，不重复抽样的抽样平均误差的计算，采用重复抽样的公式的场合是（　　）。

A．样本单位数占总体单位数的比重很小时

B．样本单位数占总体单位数的比重很大时

C．样本单位数目很少时

D．样本单位数目很多时

4．在其他条件不变的情况下，样本单位数和抽样误差的关系是（　　）。

A．样本单位数目越大，抽样误差越大

B．样本单位数目越大，抽样误差越小

C．样本单位数目的变化与抽样误差的数值无关

D．抽样误差变化程度是样本单位数目变动程度的一半

5．先将总体单位按某一标志顺序排列，然后按照固定顺序和相同间隔来抽取样本单位的抽样组织方式是（　　）。

A．简单随机抽样　　B．分类抽样　　C．等距抽样　　D．整群抽样

二、判断题

1．概率度 t 与概率保证程度之间没有确定的关系。（　）

2．在相同条件下，重复抽样的抽样误差大于不重复抽样的抽样误差。（　）

3．在相同条件下，不同的抽样组织形式却有相同的抽样误差。（　）

4．要提高估计的可靠程度，必须扩大误差范围。（　）

5．实际调查时只能按照计算的必要样本容量来抽取样本数目。（　）

三、计算题

1．某企业对 10 000 件新产品进行抽查，随机抽取 500 件，发现其中 20 件不合格。要求：以 90%的概率保证程度，推断全部新产品的合格率。

2．已知某企业职工的收入情况如表 7-3 所示。

表 7-3　某企业职工的收入情况表

不同收入类型	职工人数（人）	年平均收入（百元）	各类职工收入的标准差（百元）
较高的	200	1 320	48
一般的	1 600	804	30
较低的	1 200	600	45
合计	3 000		

要求：按照职工人数的 5%抽样，计算以下几个问题：

（1）抽样年平均收入；

（2）年平均收入的抽样平均误差；

（3）在 95%的概率保证程度下，职工月平均收入的可能范围。

3．对某城市全部 96 000 户职工家庭生活抽样调查，已知过去该市职工家庭生活收入的标准差为 40 元，允许误差不得超过 10 元，要求：在 95.45%的概率保证程度下，进行简单随机不重复抽样时，应抽取多少户？



课外延伸阅读

PPS 抽样

概率比例规模抽样（Probability-Proportional-to-Size Sampling，PPS 抽样）是一种运用属性抽样原理对货币金额而不是对发生率得出结论的统计抽样方法。

PPS 抽样是以货币单位作为抽样单元进行选样的一种方法，有时也被称为金额加权抽样、货币单位抽样、累计货币金额抽样及综合属性变量抽样等。在该方法下总体中的每个货币单位被选中的机会相同，所以总体中某一项目被选中的概率等于该项目的金额与总体金额的比。项目金额越大，被选中的概率就越大。但实际上注册会计师并不对总体中的货币单位实施检查，而是对包含被选取货币单位的余额或交易实施检查。注册会计师检查的余额或交易被称为逻辑单元或实物单元。

（1）PPS 抽样的优点之一是无须考虑被审计金额的预计变异性。

因为 PPS 抽样是按照金额来进行的，那么每一块钱被抽到的可能性都是一样的，那么 1000 元比 100 元抽样的可能性就要大，所以这也就表明 PPS 抽样本身的原理就已经考虑了抽样单元变异性的特点，分别赋予了不同金额的抽样单元不同的抽取概率。所以 PPS 抽样的样本规模不需考虑被审计金额的预计变异性。

（2）在 PPS 抽样中，如果样本被重复选择，是不影响样本量的，同时，如果发现错报，则在评价样本结果时，重复出现了几次就算几次错报。

（3）为什么 PPS 抽样使用的是属性抽样的原理？

PPS 抽样是以抽样原理为基础的，PPS 抽样是以错报的发生率来计算总体的错报金额的。

通过审计抽样的学习，属性抽样中用的是比率，变量抽样用的是错报金额。

（4）PPS 抽样的优缺点中“PPS 抽样中如果项目金额超过选样间距，PPS 系统将自动识别所有单个重大项目”如何理解？

PPS 抽样是以货币单位作为抽样单元进行抽样的一种方法，在该种计算方法下总体中的每个货币单位被选中的机会相同，所以总体中某一项目被选中的概率等于该项目的金额与总体金额的比。它有助于注册会计师将审计的重点放在较大的余额或交易上。此抽样方法之所以得名，是因为总体中每一个余额或交易被选中的概率与其账面金额成比例。因为它的特性因此其具备了其特有的优缺点。

举一个日常生活中的例子来理解这个问题：

例如某地的商品房比较抢手，有人半夜就排队了，即使这样不是说排到队的人都有机会买到房，开发商届时会抽号，从第一个人开始每隔一个随机数（比如 4）开始抽取，凡抽中的人均可购住房一套。现在某户人家共计五口人一起来排的队，这五口人是一家，这一家是一个逻辑单元，而其中的每一个人都是一个 PPS 的抽样单元，由于现在是每隔 4 个人选中一个，那么这家肯定能有一个人被选中，但最终买到房子的并不是这一人，而是以这个人为代表的这一家，上面的这个例子就可以说明为什么“PPS 抽样中如果项目金额超过选样间距，PPS 系统选样自动识别所有单个重大项目”。

资料来源：中国审计网. http://www.iaudit.cn/Article/ShowArticle.asp?ArticleID=153204

第8章 相关与回归分析

学习目标要求

1. 理解现象之间存在的相关关系，正确比较和判断变量之间的相关程度。
2. 掌握相关分析和回归分析基本理论和基本方法。
3. 能够依据实际资料对现象之间的相关关系进行分析和预测。

重点难点提示

1. 相关系数的计算与分析。
2. 线性回归方程的建立，并根据线性回归方程进行推断与预测。

教学引导案例

发改委：2016 年 1～4 月全国铁路货运量 10.5 亿吨 同比下降 7.9%

据国家发改委网站消息，统计月报数据显示，2016 年 1～4 月全国铁路完成货运量 10.5 亿吨，同比下降 7.9%，降幅比 1～3 月收窄 1.1%，完成货物周转量 7473 亿吨公里，同比下降 8.3%，降幅比 1～3 月收窄 1%。

数据还显示，4 月份，全国铁路完成货运量 2.6 亿吨，同比下降 4.5%，完成货物周转量 1859 亿吨公里，同比下降 5.2%。

1～4 月全国公路完成货运量 94.5 亿吨，同比增长 3.7%，完成货物周转量 17 422 亿吨公里，同比增长 3%。4 月份，全国公路完成货运量 28.3 亿吨，同比增长 5.1%，完成货物周转量 5172.8 亿吨公里，同比增长 3.3%。

1～4 月全国水路完成货运量 18.9 亿吨，同比增长 1.4%，完成货物周转量 28 592 亿吨公里，同比下降 1.8%。4 月份，全国水路完成货运量 5 亿吨，同比增长 0.7%，完成货物周转量 7101.8 亿吨公里，同比下降 7.3%。

1～4 月，民航客货运输均保持较快增长态势，全国民航完成货运量 204.5 万吨，同比增长 3.9%，完成货物周转量合计 67.8 亿吨公里，同比增长 5.7%。1～4 月，全国民航完成客运量 1.55 亿人，同比增长 10.8%，完成旅客周转量 2658.7 亿人公里，同比增长 14.7%。

1～4 月，全国规模以上港口完成货物吞吐量 37.7 亿吨，同比增长 2.2%，其中外贸货物

吞吐量12.1亿吨，同比增长2.8%。规模以上港口完成集装箱吞吐量6826万标准箱，增长2.1%。

资料来源：http://finance.ifeng.com/a/20160530/14441110_0.shtml

思考：

1. 案例中涉及哪些重要变量？

2. 这些变量之间的关系如何？公路等运输方式货运量的增加是否导致了铁路货运量的下降？

第一节　相关分析概述

一、相关关系的概念

客观世界中，任何事物或现象都不是孤立存在的，它总是和其他事物或现象相互联系、相互制约的。每一现象的存在和发展一方面影响着周围一些事物的存在和发展，另一方面又受周围一些事物的影响和制约。现象之间的这种联系可以通过数量对应关系反映出来，因此现象之间的联系必然表现为变量之间的依存关系。变量之间的数量关系存在着两种不同的类型：一种是函数关系；另一种是相关关系。

（一）函数关系

函数关系是指变量之间存在着严格的、确定性的依存关系。在这种关系中，对于某一变量的每一个数值，都有另一个变量完全确定的值与之相对应，并且这种关系可以用一个数学表达式反映出来。例如，圆的面积 s 与圆的半径 r 之间存在着函数关系，即 $s=\pi r^2$，当圆的半径确定以后，圆的面积也随之确定；销售收入 Y 与所销售的产品数量 X 之间存在着函数关系 $Y=PX$（P 表示单位产品的价格）。

（二）相关关系

相关关系是指客观现象之间存在不确定的数量上的关联关系，即对于某一变量的每一个数值，另一变量的值虽然不确定，但它仍按某种规律在一定范围内变化。在各种经济活动和生产过程中，许多经济的、技术的因素之间都存在着这种相关关系，分析这种关系的内在联系和表现形式是统计研究工作的一项重要任务。

相关关系具有以下两个特征。

1. 现象之间存在数量上的关联关系

相关关系表现为，当一个现象发生数量上的变化时，另一个现象也会相应地发生数量上的变化。例如，随着家庭收入的增加，会相应带来储蓄存款的增加；经济的增长，会带来用电量的增加；公司广告费支出的增加，会影响销售额的增长；商品价格的高低变动，会影响其销售量的大小等。

但是相关关系不同于因果关系，如人的身高和体重，不太好区分谁是原因谁是结果。

2. 现象之间的数量依存关系是不确定的

在相关关系中，虽然两个变量的数量间有依存关系，但其数值是不固定的。例如，身高相同的人，体重却会有许多不同的值；同样的施肥量，单位面积产量则可能出现不同的数值；同样收入水平的家庭，银行存款额也不会相同；等等。之所以会发生这种不确定性的情况，是因为还受到许多其他因素的影响。

二、相关关系的种类

现象间的相关关系可以按不同的标志分类，不同的相关关系往往要用不同的方法进行研究。

（一）按影响因素的多少不同分类

按影响因素的多少不同，相关关系可分为单相关、复相关和偏相关。

1. 单相关

两个变量的相关，即一个变量对另一变量的相关关系称为单相关或简单相关。例如，在计件工资的条件下，快递员的月工资只与其完成递送量成单相关关系。

2. 复相关

三个或三个以上变量的相关关系称为复相关，又称多元相关。例如，物流成本与物流运输里程、员工工资及管理费用之间的关系是复相关关系。

3. 偏相关

就多个变量测量两个变量的相关程度而假定其他变量不变者，称为偏相关。例如，物流公司的运输总成本与货物运输量和运输里程的关系是复相关，如果假定货物运输量不变，运输总成本和运输里程的关系是偏相关。

在实际工作中，如果存在多个因素对现象的影响，应该加以筛选，抓住其中最主要的因素，研究其相关关系。

（二）按相关的形式不同分类

按相关的形式不同，相关关系可分为线性相关和非线性相关。

对两个具有相关关系的现象进行实际调查，获得一系列成对的数据后，根据两变量的各对相应数值，在平面直角坐标系中描出的若干个点，如果散点趋于一条直线，则称为线性相关（如图 8-1 和图 8-2 所示）；若散点趋向于某种曲线，则称为非线性相关或曲线相关（如图 8-3 所示）。

例如，货物运输量与企业总产值之间，大致上成线性相关，而商品流转额和流通费用率之间通常是非线性相关。

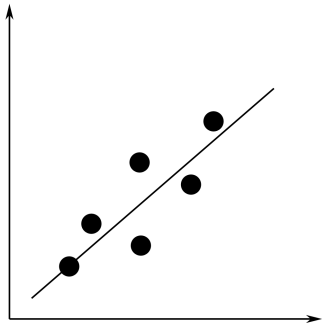


图 8-1　正线性相关

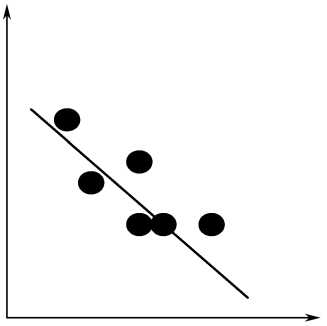


图 8-2　负线性相关

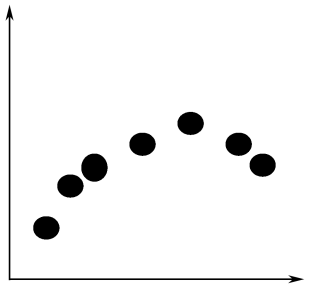


图 8-3　非线性相关

（三）按相关的方向不同分类

按相关的方向不同，线性相关又分为正相关和负相关。

当一个变量的数值增加或减少时，另一变量的数值也相应地增加或减少，即二者做同方向的变化，称为正相关（如图 8-1 所示）。反之，当一个变量的数值增加，另一变量的数值反而减少；或一个变量的数值减少，另一变量的数值反而增加时，即二者做反方向的变化，则称为负相关（如图 8-2 所示）。例如，商品流转的规模越大，而流通费用水平则越低。必须指出，许多现象正负相关的关系仅在一定范围内存在。

对于非线性相关的方向，一般不进行区分。

（四）按相关的程度不同分类

按相关的程度不同，相关关系可分为完全相关、不相关、不完全相关。

1. 完全相关

一个变量的数量变化由另一个变量的数量变化所确定，则称这两个变量间的关系为完全相关。在这种情况下相关关系即为函数关系，可以用相应的方程式来准确地表示。根据两变量的各对相应数值，在平面直角坐标系中描出的若干个点，构成一条直线。例如，匀速行驶的车辆的路程（s）取决于车辆行驶的时间（t），即 $s=vt$，如图 8-4 和图 8-5 所示。

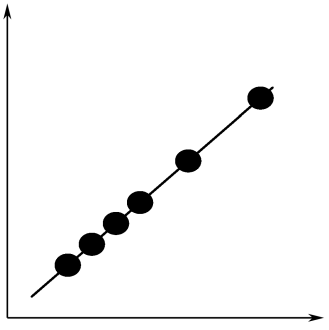


图 8-4　完全正相关

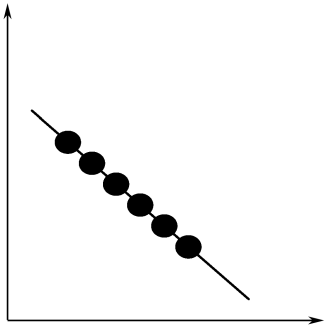


图 8-5　完全负相关

2. 不相关

若两个变量彼此互不影响，其数量变化各自独立，则称为不相关或零相关。根据两变量

的各对相应数值，在平面直角坐标系中描出的若干个点散点分布趋近球形就属于不相关（如图 8-6 所示）。例如，运输车辆的载货量与员工出勤率分属不同总体的现象，一般认为是不相关的。

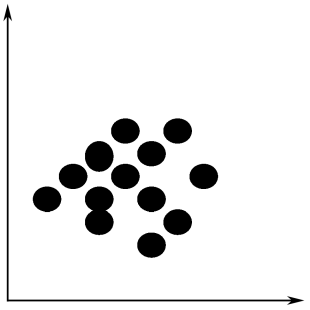


图 8-6　不相关

3. 不完全相关

一个变量的数量变化会随另一个变量的变化而变化，但是两变量之间没有函数关系，介于完全相关和不相关之间，则称为不完全相关，这是相关分析的主要研究对象，如图 8-1 和图 8-2 所示。

三、相关分析的含义、内容与意义

（一）相关分析的含义

相关分析是研究两个或两个以上变量之间的相关方向和相关密切程度的统计分析方法。在相关分析中不必确定变量中哪个是自变量、哪个是因变量。其所涉及的变量都是随机变量，是对等关系。

（二）相关分析的内容

1. 确定变量之间有无相关关系及相关关系的表现形式

这是相关分析中的定性分析。确定相关关系是否存在是进行相关分析的前提，它包括确定变量之间是否有相关关系，是否真实相关。在通过分析确定有真实相关关系之后，要进一步确定其表现形式是直线相关还是曲线相关。相关形式不同，分析方法也不同。

2. 确定变量之间相关的密切程度

若只是粗略地直观认识变量间的密切程度，则要编制相关表或绘制相关图。若要定量地确定变量之间相关关系的密切程度，则要计算其相关系数。如果变量之间有相关关系，但关系不密切，进一步研究就没有多大价值。

3. 建立合适的数学模型

根据相关关系的表现形式，即建立变量之间数量变化关系的近似表达式，需要配合相应

的数学模型。若变量之间的关系表现为线性相关，就采用线性方程拟合；若变量之间表现为曲线相关，就采用曲线方程拟合。以此来表现变量之间相互依存关系数量上的规律性，从而为推算、预测提供依据。

4. 测定变量估计值的可靠程度

配合直线或曲线方程，将自变量数值代入方程中就可以估计或预测因变量的值。估计值与实际观察值一般是有差别的，差别的大小可以通过计算估计标准误差来确定。估计标准误差小，说明估计或预测值较准确，可靠程度较高；估计标准误差大，则估计或预测值不够准确，可靠程度低。

（三）相关分析的意义

任何社会经济现象的存在和运动都或多或少地受到其他事物或现象的影响和制约，它们彼此依赖，相互联系。例如，企业通过增加广告次数和频率增强对消费者的刺激，从而增加商品的销量。相关分析是为了揭示现象之间的联系方向和联系密切程度的一种统计分析方法；回归分析则是在相关分析的基础上，进一步揭示一事物影响另一事物变动的一般水平，以便从一个已知量来推算或预测另一个未知量的一种统计分析方法。

例如，实验数据的一般处理、经验公式的求得、因素分析、产品质量的控制、气象及地震预报、自动控制中数学模型的制定等，这类统计分析方法可以帮助人们提高分析问题、预测事物发展变化的精确性和判断、决策的科学性，因而在社会经济活动中得到了广泛运用。

第二节　简单线性相关分析

一、相关关系的判断

进行相关分析时，首先要对现象之间是否存在相关关系，以及具有什么样的相关关系进行分析，做出判断。分析判断的方法有两种。

（一）根据对客观现象的定性认识来判断

现象之间有无关系、有什么样的关系，这是一种质的规定性。对于这种质的规定性的认识属于定性认识。从认识的一般规律来讲，只有在定性判断的基础上，才能进行定量的分析和判断。定性认识是否准确取决于研究者的理论知识、专业知识、实际经验和分析研究能力。所以，对现象进行定性判断是进行相关分析的基础和前提。

（二）利用相关表和相关图进行判断

在研究者对社会经济现象之间是否存在相关关系及相关关系的类型不易做出准确的判断时，还可以根据掌握的统计资料，编制相关表和绘制相关图来做出进一步的判断。相关图表有助于直观地判断现象之间有无联系，可以观察现象相关关系的类型和密切程度。

1. 相关表

相关表是根据现象变动的实际数据资料编制出来的反映两个变量之间相关关系的统计表。

简单相关表是资料未经分组的相关表，它是把一个变量的取值按照从小到大的顺序并配合另一变量的取值一一对应而平行排列起来的统计表。简单相关表是现象变量之间相关研究初步结果的表现。例如，研究某物流公司产量和分拣成本之间的关系时，收集到 2016 年 1～12 月的资料如表 8-1 所示。

表 8-1　产量和分拣成本的资料

时　期	产量 x（件）	分拣成本 y（元）
1 月	250	275
2 月	200	240
3 月	300	315
4 月	360	350
5 月	400	388
6 月	350	360
7 月	380	385
8 月	320	332
9 月	315	320
10 月	308	315
11 月	298	310
12 月	280	290

从表 8-1 可以看出，产量与分拣成本有同步增长的趋势，两变量存在着正相关关系。简单相关表适用于总体单位数相对较少的原始资料。

2. 相关图

通过编制相关图也可直观地判断现象之间大致上呈现何种相关关系。相关图又称散点图或相关散布图，它是以直角坐标系的横轴代表自变量 x，纵轴代表因变量 y，并将两变量相对应的变量值用坐标点形式描绘出来，用以反映相关点分布状况的图形。

相关图上每一个点代表一对观测值。通过相关图可以较为形象直观地判断两个变量之间有无相关关系、相关关系的类型和密切程度。对于表 8-1，由于已假设产量为自变量，分拣成本为因变量，则产量作为 x 轴，而分拣成本作为 y 轴。现将数据在图中用点描出来，作出相关图，如图 8-7 所示。

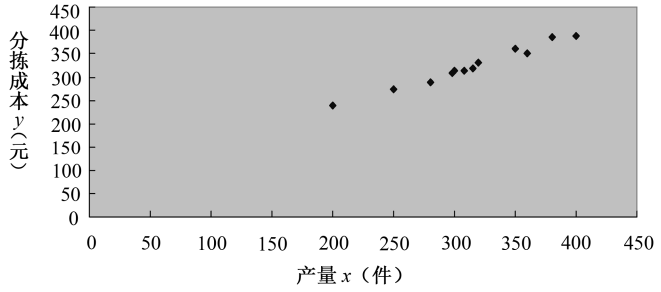


图 8-7　产量和分拣成本相关图

从图 8-7 可以看出，产量与分拣成本的关系虽然不十分严格，但有直线相关的趋势，而且大致可以看出关系比较密切。从相关图可以判断产量与分拣成本这两个变量是相关的。另外，从这些散点的分布状况看，这两个变量是正相关的，也就是说，随着产量的增加，分拣成本也有所增大。所以，相关图比相关表更明显、直观地表现出现象之间的相关关系。

二、相关系数

（一）相关系数的概念

相关表和相关图只能粗略地反映现象之间的相关关系，但不可能准确判断相关密切程度。要判断现象之间相关关系的密切程度，必须计算相关系数。相关系数是指在直线相关条件下，反映两个现象之间相关关系密切程度的统计分析指标，用 r 表示。

（二）相关系数的计算

相关系数的计算方法有若干种，最易理解的一种叫积差法，直接来源于数理统计中相关系数的定义。这里我们不做公式的推导和证明，只简单介绍其结果。

相关系数的定义公式为：

$$r=\frac{\sigma^2_{xy}}{\sigma_x\sigma_y}$$

式中：r 代表相关系数；

$\sigma^2_{xy}=\dfrac{\sum(x-\bar{x})(y-\bar{y})}{n}$ 代表变量 x 与 y 的协方差；

$\sigma_x=\sqrt{\dfrac{\sum(x-\bar{x})^2}{n}}$ 代表变量 x 的标准差；

$\sigma_y=\sqrt{\dfrac{\sum(y-\bar{y})^2}{n}}$ 代表变量 y 的标准差。

根据 σ^2_{xy}、σ_x、σ_y 的表达式，经代数演算，相关系数的计算公式又可表述为：

$$r=\frac{n\sum xy-\sum x\sum y}{\sqrt{n\sum x^2-(\sum x)^2}\sqrt{n\sum y^2-(\sum y)^2}}$$



【例 8-1】

某医院在 2016 年对医用设备的使用年限和维修费支出情况进行了调查，调查结果如表 8-2 所示。

表 8-2　设备使用年限与维修费资料表

序　号	设备使用年限（年）	年维修费（元）	序　号	设备使用年限（年）	年维修费（元）
1	2	400	7	5	800
2	2	540	8	6	700
3	3	520	9	6	760
4	4	640	10	6	900
5	4	740	11	8	840
6	5	600	12	9	1080

初步判断这两个变量是相关的。根据调查资料，计算医用设备的使用年限与年维修费用的相关系数如表 8-3 所示。

表 8-3　相关系数计算表

序　号	使用年限 x（年）	年维修费用 y（元）	x^2	y^2	xy
1	2	400	4	160 000	800
2	2	540	4	291 600	1080
3	3	520	9	270 400	1560
4	4	640	16	409 600	2560
5	4	740	16	547 600	2960
6	5	600	25	360 000	3000
7	5	800	25	640 000	4000
8	6	700	36	490 000	4200
9	6	760	36	577 600	4560
10	6	900	36	810 000	5400
11	8	840	64	705 600	6720
12	9	1080	81	1 166 400	9720
合　计	60	8520	352	6 428 800	46 560

根据计算表和相关系数的计算公式，计算相关系数得出：

$$r=\frac{n\sum xy-\sum x\sum y}{\sqrt{n\sum x^2-(\sum x)^2}\sqrt{n\sum y^2-(\sum y)^2}}$$

$$=\frac{12\times 46\,560-60\times 8\,520}{\sqrt{12\times 352-(60)^2}\times\sqrt{12\times 6\,428\,800-(8\,520)^2}}$$

$$=0.8913$$

（三）相关关系的分析

计算相关系数的目的是判断两个变量之间的线性相关程度与方向，因此，需要明确判断标准。相关系数 r 的性质可归纳如下：

①相关系数 r 的取值范围永远在-1 到+1 之间，即 $-1 \leqslant r \leqslant 1$。

②当 $-1 < r < 0$ 时，相关图呈现出 y 随 x 的增加而减少的趋势，即为负相关关系。

③当 $r=0$ 时，表示 x 与 y 两个变量之间不存在线性相关关系。但需要注意的是，$r=0$ 只能说明两个变量之间没有线性关系，而不能排斥其他关系，如可能存在非线性相关关系。变量之间的非线性相关程度较大时，可能会导致 $r=0$。因此，当 $r=0$ 或很小时，应结合相关图做出合理的解释，而不能轻易得出两个变量之间不存在相关关系的结论。

④当 $0 < r < 1$ 时，相关图呈现出 y 随 x 的增加而增加的趋势，即为正相关关系。

⑤当 $|r|=1$ 时，相关图呈现出一条直线，表示变量 x 与 y 完全线性相关，实际上这时两变量之间存在着确定的函数关系。

为了判断现象之间相关程度的高低，一般将相关关系密切程度划分为四个等级：$|r| \leqslant 0.3$ 称为微弱相关；$0.3 < |r| \leqslant 0.5$ 称为低度相关；$0.5 < |r| \leqslant 0.8$ 称为显著相关；$0.8 < |r| < 1$ 称为高度相关。

对于例 8-1 中设备的使用年限与维修费用的相关系数 0.8913，可以判断：设备使用年限与年维修费用之间存在着高度的正相关，即设备使用年限越长，年维修费用越高。

三、线性相关分析的特点

线性相关分析的特点是：①相关系数是测定变量之间相关密切程度和相关方向的代表性指标，在分析两个变量的直线相关关系时，通常计算直线相关系数；②参与相关分析的两个变量是对等的，不分自变量和因变量，因此相关系数只有一个；③相关系数有正负号，反映相关关系的方向，正号反映两变量呈正相关关系，负号反映两变量呈负相关关系；④相关分析的两个变量都是随机变量。

计算相关系数，在相关分析和经济预测中都有很重要的用途。回归分析与相关系数的计算有着紧密的关系，一般来说，只有显著相关时，进一步作回归分析才有意义。

第三节　回 归 分 析

一、回归分析的概念、类型和特点

（一）回归分析的概念

通过相关分析，可以判定变量之间是否存在相关关系及相关的密切程度，但它不能根据

一个变量的值估计推算出另一个变量的值。也就是说，它不能说明两个变量之间的一般数量关系。而回归分析的基本思想恰恰是：虽然自变量和因变量之间没有严格的、确定性的函数关系，但可以设法找出最能代表它们之间关系的数学表达形式。

回归分析是对具有相关关系的多个变量之间的数量变化进行数量测定，配合一定的数学方程，以便对因变量估计或预测的一种统计分析方法。

根据回归分析的方法得出的数学表达式称为回归方程或回归模型。回归方程有多种形式，可以是直线回归方程，也可以是曲线回归方程。用回归方程来表明两个变量之间线性相关关系的方程式称为简单线性回归方程或一元线性回归方程，这种分析方法称为一元线性回归分析方法。

在回归分析中，变量之间存在一定的数量关系，但又不呈现函数关系，即观察值不全落在回归线上，而是散布在回归线周围。但离回归线越近观察值越多，偏离较远的观察值极少。

（二）回归分析的类型

1. 一元回归和多元回归

根据自变量的个数不同，可将回归分析分为一元回归分析和多元回归分析。只有一个自变量的回归分析称为一元回归分析，又称简单回归。有两个或两个以上自变量的回归分析称为多元回归分析，或称复回归。

2. 线性回归和非线性回归

根据回归线的形状不同，可将回归分析分为线性回归分析和非线性回归分析。线性回归分析是指变量之间关系的形态呈直线趋势，故而可用直线方程来描述它们之间的数量关系。若变量之间是非线形相关关系，可通过建立非线形回归方程来反映它们之间的数量关系，即非线性回归分析。

将以上两种分类方式结合起来，就会有一元线性回归分析和多元线性回归分析、一元非线性回归分析和多元非线性回归分析。本章将重点介绍一元线性回归分析。

（三）回归分析的特点

回归分析具有以下五个特点：①两个变量中，一个是自变量，另一个是因变量：②回归方程不是抽象的数学模型，而是用自变量数值推算因变量数值的根据，必须反映变量之间关系的一般变动情况；③对于没有明显因果关系的两个变量，可以确定两个不能互相替代的回归方程，一个是以 x 为自变量、以 y 为因变量的回归直线方程，另一个是以 y 为自变量、以 x 为因变量的回归直线方程，这两条回归直线方程斜率不同，意义也不同；④一元线性回归方程系数（即斜率）有正有负，正回归系数表明两变量之间正相关，负回归系数表明两变量之间负相关，至于回归系数的大小，视原数列使用的计算单位而定，这不能表明两变量之间的变动程度；⑤计算回归方程的资料要求是，因变量为随机的，而自变量是给定的数值，求出回归方程后，将给定自变量值代入方程中，推算出因变量的一般值或平均数值。

二、回归分析的基本内容

1. 确定相关关系的数学表达式

为了测定相关现象之间数量变化的一般关系，可以建立函数关系的近似表达式即回归方程作为相关关系的数学表达式。如果现象之间表现为直线相关，则可以用配合直线方程的方法；如果表现为曲线相关，则采用配合曲线方程的方法。这是进行判断、推算、预测的依据。

2. 依据回归方程进行回归估计或预测

由于回归方程反映了变量之间的一般性数量关系，因此当自变量数值发生变化后，可依据回归方程估计出因变量可能发生相应变化的数值。因变量的回归估计值虽然不是一个必然的对应值，但它至少可以从一般性角度或平均意义角度反映因变量可能发生的数量变化。

3. 确定因变量估计值误差的程度

用配合直线或曲线方程的方法，可以求得反映变量之间数量变化的关系式，从而计算出许多因变量的估计值。但是，根据这个直线或曲线方程而得到的因变量的估计值与实际值是有差异的。这就要计算反映因变量估计值与观察值之间差异程度的指标——估计标准误差。差异小表示估计值比较准确，差异大说明估计值不够准确。作为统计预测需要掌握这一准确度的大小。

三、回归分析与相关分析的关系

回归分析和相关分析是研究变量之间相互联系的两种统计方法，对两个变量间的关系进行完整的分析应该包含两者。它们既有密切的联系又有区别。

（一）联系

1. 相关分析是回归分析的基础和前提

如果缺少相关分析，没有从定性上证明现象间是否具有相关关系，没有对相关关系的密切程度做出判断，就不能进行回归分析。即使勉强进行了回归分析，也毫无意义。

2. 回归分析是相关分析的深入和继续

相关分析需要回归分析来表明现象数量关系的具体形式。因此，只有进行回归分析，拟合了回归方程，才可能进行有关分析和预测，相关分析才有实际意义。

（二）区别

①相关分析所研究的两个变量有对等关系，即不必确定两个变量中哪个是自变量、哪个是因变量，改变两变量的地位并不影响相关系数的数值。回归分析所研究的两个变量之间的关系是不对等的，必须根据研究的目的确定哪个是自变量、哪个是因变量。

②对两个变量 x 和 y 来说，相关分析只能计算出一个反映两个变量之间相关密切程度的

相关系数，计算中改变 x 和 y 的地位不影响相关系数的数值。回归分析可根据研究目的的不同分别建立两个不同的回归方程：一个以 x 为自变量、y 为因变量，可以得出 y 对 x 的回归方程；另一个以 y 为自变量、x 为因变量，可以得出 x 对 y 的回归方程。若要画出图来，则是两条斜率不同的回归直线。

③相关分析计算的相关系数是一个绝对值在 0～1 之间的抽象系数，其数值大小反映变量之间相关关系的程度。回归分析建立的回归方程反映的是变量之间变动的内在联系和比例关系，不是抽象系数。根据回归方程，利用自变量的给定值，可以估计或推算因变量的数值，估计出来的参数都是有实际经济含义的数值。

④相关分析对资料的要求是,两个变量都必须是随机变量，各自接受随机因素的影响。而回归分析对资料的要求是,因变量是随机的，而自变量则可以不是随机的，是给定的数值。

四、一元线性回归分析

（一）建立回归方程

一元线性回归分析又称简单线性回归分析。它是一种对具有显著直线相关的两个变量间数量变化进行分析，确定回归方程，以预测估计因变量数值的方法。分析时所建立的回归方程称为一元线性回归方程，也称直线方程。它是分析一个自变量 x 与一个因变量 y 之间线性关系的数学方程。其回归方程的基本形式是：

$$y_c = a + bx$$

式中：y_c 是因变量 y 的估计值，或称回归值、理论值；

a 是直线与纵轴的交点，数学上称为截距，表示当自变量等于零时因变量的起点值；

b 是直线与横轴间夹角的正切值，数学上称为斜率，回归分析中称为回归系数，表示自变量 x 每变动一个单位时因变量 y 的平均变动量。当 b 的符号为正时，自变量和因变量按相同方向变动；当 b 的符号为负时，自变量和因变量按相反方向变动。

（二）确定回归方程中的参数

a 和 b 是回归方程中的两个待定参数，而它们一旦被确定，这条直线便被唯一确定了。但用于描述这 n 组数据的直线有许多条，究竟用哪条直线来代表两个变量之间的关系，则需要有一个明确的原则。必须选择一条距离各散点最近的直线来代表变量 x 与 y 之间的关系。根据这一思想确定直线中未知常数 a、b 的方法称为最小平方法（最小二乘法）。利用这种方法得到的直线方程，可以使得推算的估计值 y_c 与实际值 y 的离差平方和达到最小。用公式表示为：

$$\sum (y - y_c)^2 = \sum (y - a - bx)^2 = \text{最小值}$$

设 $Q = \sum \left(y - y_c\right)^2$，则 Q 是两个待定参数 a 和 b 的函数。根据微积分学中求极值的原理，需分别对 a 和 b 求偏导数，并令其等于零。经过整理，得出由下列两个方程式组成的标准方程组：

$$\begin{cases} \sum y = na + b\sum x \\ \sum xy = a\sum x + b\sum x^2 \end{cases}$$

解这个方程组得：

$$b = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sum (x-\bar{x})^2} = \frac{n\sum xy - \sum x\sum y}{n\sum x^2 - (\sum x)^2}$$

$$a = \frac{\sum y}{n} - b\frac{\sum x}{n} = \bar{y} - b\bar{x}$$

a、b的值确定后，即可求出y关于x的回归方程$y_c = a + bx$。

配合线性回归方程的前提条件是两个变量之间确实存在显著的直线相关关系；否则配合直线回归方程将毫无意义。因此，在进行回归分析之前，可以通过相关图及计算相关系数的方法，在确定其相关程度显著的条件下，再配合直线回归方程。

【例 8-2】

根据表 8-4 中的资料，建立医用设备的使用年限与年维修费用的线性回归方程。

通过例 8-1 计算的相关系数 0.8913，已知设备使用年限x和年维修费用y这两个变量有高度的线性相关关系。虽然各散点并不在一条直线上，但可配合一条直线，近似地表达现象变动的一般规律。

将表 8-5 中的资料代入参数a和b的求值公式，得到：

$$b = \frac{n\sum xy - \sum x\sum y}{n\sum x^2 - (\sum x)^2} = \frac{12\times 46\,560 - 60\times 8520}{12\times 352 - 60^2} \approx 76.15$$

$$a = \frac{\sum y}{n} - b\frac{\sum x}{n} = \frac{8520}{12} - 76.15\times\frac{60}{12} \approx 329.25$$

将a和b的数值代入方程式$y_c = a + bx$，求得直线回归方程为：

$$y_c = 329.23 + 76.15x$$

在回归方程中，a=329.25 是回归直线在y轴上的截距，表示年维修费用的理论起点值；回归系数b=76.15，表示设备使用年限每增加一年，年维修费用平均增加 76.15 元。

（三）利用回归方程进行预测

利用所求的回归方程，只需把自变量的值代入上述回归直线回归方程，就可以得到对应的因变量的预测值。在例 8-2 中，当使用年限 x=8 年时，年维修费用的估计值 $y_c = 329.25 + 76.15\times 8 = 938.45$（元）。

通过回归方程不仅可以推出已知值的估计值，而且可以预测超出现有数据范围的未知值。例如，当使用年限$x = 10$年时，维修费用的估计值为$y_c = 329.25 + 76.15\times 10 = 1090.75$（元）。需要注意的是，预测超出现有数据的范围有时是靠不住的。

例如，施肥量和农作物生产量只在一定范围内才具有正相关关系。施肥量超过一定限度，产量不但不会增加，反而会减少，其他许多现象也是如此。因此用相关分析和回归方程分析方法进行推算和预测时要注意它的作用范围。

五、估计标准误差

（一）估计标准误差的概念

建立回归方程的目的主要是推算或预测，即要根据给定的自变量的数值来推算未知的因变量的数值。但推算出来的因变量的数值只是一个可能的估计值，代表一般关系的平均值，它和实际值是有出入的。因为在研究社会经济现象的变动时，不可能把影响现象变动的各种因素都考虑到，所以，因变量的每一个实际值 y 和与其对应的估计值 y_c 并不完全相等。因此，需要对估计值的代表性进行评价。通常采用计算估计标准误差的方法。

这个推算结果的准确程度如何，主要取决于所配合的回归直线的代表性。故回归直线的代表性和推算结果的准确性是一个问题的两个方面。在进行回归分析时，必须弄清这个问题。

估计标准误差是指因变量的实际值 y 与估计值 y_c 的平均离差，估计标准误差是说明使用回归方程推算结果的准确程度的分析指标，也是反映回归直线代表性大小的分析指标。估计标准误差和标准差的性质相同，都是说明离散程度的指标。

（二）估计标准误差的作用

1. 说明以回归直线为中心的所有相关点的离散程度

估计标准误差数值大，则说明平均误差越大，相关点与回归直线的离散程度越大；反之则说明离散程度小。这个数值的大小反映了利用回归直线进行估计或预测的准确程度。

2. 说明回归直线的代表性大小

估计标准误差大，则回归直线的代表性小，它的实用价值也小；估计标准误差小，则回归直线的代表性大，它的实用价值也大。

从上述作用来看，估计标准误差和相关系数一样，也具有说明相关关系密切程度的作用。不同的是相关关系数越大越好，估计标准误差越小越好。相关系数用相对数表现，密切程度的概念比较明确；估计标准误差用绝对数表现，关系密切的程度表示得不那么明显，它也不能说明是正相关还是负相关。

3. 在抽样调查条件下，估计标准误差是计算回归抽样误差的一个根据

就像总体方差是计算平均指标抽样误差的根据一样，计算回归抽样误差时应该使用总体的估计标准误差。但总体的信息常常是没有的，要用样本的估计标准误差来代替。

（三）估计标准误差的计算

估计标准误差的计算原理与标准差基本相同，其定义公式为：

$$S_y=\sqrt{\frac{\sum(y-y_c)^2}{n-2}}$$

式中，S_y 代表估计标准误差；y 代表因变量实际观察；y_c 代表根据回归方程推算出的因变量的估计值；n−2 代表自由度。

S_y 数值的大小说明回归估计值的准确程度和回归方程的代表性。S_y 值越大，说明估计

值的准确程度越低，回归方程的代表性越小；S_y值越小，说明估计值的准确程度越高，回归方程的代表性越大。若S_y=0，则说明实际观察值 y 与估计值y_c没有差异，也就是说变量之间的关系就是回归方程所表现出来的那种函数关系，从相关图中可以显示实际观察值全部落在直线上。

下面仍以表 8-2 所示的资料及其所建立的设备使用年限和维修费用的回归方程，通过列表来说明估计标准误差的计算过程，如表 8-4 所示。

表 8-4　估计标准误差计算表

序号	使用年限 x	维修费用 y	y_c	$y-y_c$	$(y-y_c)^2$
1	2	400	481.55	−81.55	6 650.4025
2	2	540	481.55	58.45	3 416.4025
3	3	520	557.7	−37.7	1 421.29
4	4	640	633.85	6.15	37.8225
5	4	740	633.85	106.15	11 267.8225
6	5	600	710	−110	12 100
7	5	800	710	90	8100
8	6	700	786.15	−86.15	7 421.8225
9	6	760	786.15	−26.15	683.8225
10	6	900	786.15	113.85	12 961.8225
11	8	840	938.45	−98.45	9692.4025
12	9	1080	1 014.6	65.4	4277.16
合计	60	8520	8 520	—	78 030.77

将表 8-4 中的有关计算结果代入估计标准误差的计算公式，得到：

$$S_y=\sqrt{\frac{\sum(y-y_c)^2}{n-2}}=\sqrt{\frac{78\,030.77}{12-2}}=88.335$$

计算结果表明，维修费用的实际值与估计值是有差距的，平均差距为 88.335 元。为了判断估计标准误差所表明的平均离差的大小，可计算离散系数：

$$v=\frac{S_y}{\bar{y}}=\frac{88.335}{710}=0.124\,4$$

这表明回归直线与所给点列(x,y)间的配合精度相当高，以回归值y_c来估计相应的实际值y误差相当小。只有在估计标准误差较小的情况下，用回归方程进行估计或预测才具有实用价值。

实际工作中，使用定义公式计算估计标准误差计算比较麻烦。为了直接利用计算相关系数的有关数据，根据已经建立的回归方程$y_c=a+bx$，可采用简捷法计算估计标准误差。其

计算公式为：

$$s_y=\sqrt{\frac{\sum y^2-a\sum y-b\sum xy}{n-2}}$$

下面根据表 8-5 中的资料及其所建立的设备使用年限和维修费用的回归方程，用简捷法计算的估计标准误差为：

$$s_y=\sqrt{\frac{6\,428\,800-329.25\times 8\,520-76.15\times 46\,560}{12-2}}\approx 88.344$$

用定义公式和简捷公式计算估计标准误差，从理论上讲，其计算结果应该是相等的。两个式子的计算结果相差 0.009 元，这是由参数值在计算过程中四舍五入所引起的，对分析问题影响不大。



本章小结

相关分析的目的是揭示现象之间的联系方向和联系程度；回归分析则是在相关分析的基础上，进一步揭示一事物影响另一事物变动的一般水平。根据回归分析方法得出的数学表达式称为回归方程。

估计标准误差是衡量因变量的估计值与实际观测值之间平均误差大小的指标，利用此指标可以说明回归方程的代表性。估计标准误差值越大说明估计值的准确程度越低，回归方程的代表性越小；反之，说明估计值的准确程度越高，回归方程的代表性越大。



复习思考题

1．什么是相关关系？什么是函数关系？二者有什么不同？

2．什么是正相关、负相关？试举例说明。

3．绘制相关表和相关图有何作用？

4．相关分析的主要内容是什么？

5．相关系数 r 的意义是什么？怎样利用相关系数 r 来判断现象之间的相关关系？

6．什么是估计标准误差？它有什么作用？



实训基础训练

一、单项选择题

1．能够测定变量之间相关关系密切程度的主要方法是（　　）。

A．相关表　　B．相关图　　C．相关系数　　D．定性分析

2．在相关分析中，若一个变量递增时另一个变量递减，则两个变量间的关系是（　　）。

A．正相关　　B．负相关　　C．不相关　　D．复相关

3．相关系数 $|r|$ 越接近于 0 时，表示 x 与 y 的相关程度越（　　）。

A．高　　B．低　　C．不变　　D．无法确定

4．$0.5<|r|\leqslant 0.8$ 称为（　　）。

A．低度相关　　B．微弱相关　　C．显著相关　　D．高度相关

5．回归分析所研究的两个变量之间的关系是（　　），必须根据研究的目的确定哪个是自变量、哪个是因变量。

A．对等的　　B．不对等的　　C．无法确定　　D．正相关的

二、判断题

1．相关关系是指变量之间存在着严格的、确定性的依存关系。（　）

2．估计标准误差是指因变量的实际值 y 与估计值 y_c 的平均离差。（　）

3．两个相关变量的数量按相同方向变化，即相互影响的变量同时递增或同时递减，这种相关关系称为负相关。（　）

4．回归分析所研究的两个变量是对等关系，即不必确定两个变量中哪个是自变量、哪个是因变量。（　）

5．相关分析是回归分析的基础和前提。如果缺少相关分析，就不能进行回归分析。即使勉强进行了回归分析，也是毫无意义的。（　）

三、计算题

1．某企业上半年产品产量与单位成本资料如表 8-5 所示。

表 8-5　某企业产品产量和单位成本的资料表

月　份	产量（千件）	单位成本（元/件）
1	2	73
2	3	72
3	4	71
4	3	73
5	4	69
6	5	68

要求：（1）根据资料确定回归方程，计算产量每增加 1 000 件时单位成本平均变动量？

（2）当产量为 6 000 件时，单位成本为多少？

2．考察某种化工原料在水中的溶解度与温度的关系，共作了 9 组试验。已知溶解度与温度线性相关，温度为自变量，溶解度为因变量，其数据如表 8-6 所示。

表 8-6　某种化工原料在水中的溶解度与温度资料表

温度（摄氏度）	0	10	20	30	40	50	60	70	80
溶解度（克）	14	17.5	21.2	26.1	29.2	33.3	40	48	54.8

要求：（1）根据资料确定回归方程；

（2）温度每上升 1 摄氏度，溶解度将增加多少克？



课外延伸阅读

大数据：从因果分析到相关性分析

大数据已经不再是计算、统计学科的专宠，商学院开始广泛应用表明大数据正式进入各行业。统计学家纳特·西尔弗在著名的《信号和噪声》一书中说：“大数据中大多数都是不相干的噪声。除非有很好的技术信息进行过滤和处理，否则将惹上麻烦。”

杜克大学富卡商学院今年秋季开始招收大数据商业分析方向的硕士生，西安交通大学管理学院也将录取海外大数据分析的博士生作为新录取教职人员的重点之一。

大数据为我们提供了观察世界的新方式，但它往往还是类似原油粗糙的形式，没有商学院的提炼与应用，它就无法变成汽油、胶黏剂、阿司匹林、唇膏等各种现代工业产品。对于大数据来说，我们今天的时代，就像是德州刚发现油田的时代，它在信息时代的广泛应用与消费，需要各个学科的通力协作、更换思维，正如石油的发现催生了工业时代的能源革命一样。

在“前信息时代”，商学院分析消费者行为、市场结构、竞争动态、组织行为、供应链管理时，都局限于有限的样本。因为收集消费者、员工、股票、工厂等的数据都非常耗时，需要承担各种成本。即使像 IBM 这样的巨型公司，有能力将《人民日报》历年的文本输入电脑，试图破译中文的语言结构，如实现中文的语音输入或中英互译，这项技术在 20 世纪 90 年代就取得了突破，但进展缓慢，在应用中还有很多问题。

谷歌采取了不同的方法进入这个市场，它不依赖高品质的翻译，而是利用更多的数据。这家搜索巨头收集各种企业网站的翻译、欧盟的每一种语言的文本、巨大的图书扫描项目中的翻译文件。超越 IBM 以百万级的文本分析，谷歌的大数据是以十亿万级计的。其结果是，

它的翻译质量优于 IBM，能涵盖 65 种语言，而且翻译质量在云端不断优化。谷歌凌乱的大数据战胜了 IBM 少量的干净数据。

那怎样将凌乱的大数据进行对石油一样的提炼与应用呢?一项重要的思维转换就是从传统的因果分析向相关性分析转换。在传统的统计分析中，一个重要的因素是因果关系的可靠性，在有限的样本下，科学家在假设检验中往往用各种专业统计软件进行假设检验，根据概率 P 值(P-Value, Probability)进行检验决策。P 值反映某一事件发生的可能性大小，一般以 $P<0.05$ 为显著，从而确认两个变量间可能存在因果关系。

但大数据的出现改变了这种在科学界普遍追求的因果关系的检验。大数据主要从相关性着手，而不是因果关系，这从本质上改变了传统数据的开采模式。例如 2009 年 2 月，谷歌的研究人员在《自然》上发表了一篇论文，预测季节性流感的暴发，在医疗保健界引起了轰动。谷歌对 2003 年和 2008 年间的 5 000 万个最常搜索的词条进行大数据“训练”，试图发现某些搜索词条的地理位置是否与美国流感疾病预防和控制中心的数据相关。疾病预防控制中心往往跟踪全国各地的医院和诊所病人，但它发布的信息往往会滞后 1～2 个星期，谷歌的大数据却能发现实时的趋势。

谷歌并没有直接推断哪些查询词条是最好的指标。相反，为了测试这些检索词条，谷歌总共处理了 4.5 亿个不同的数字模型，将得出的预测与 2007 年和 2008 年疾病预防控制中心记录的实际流感病例进行对比后，谷歌公司发现，它们的大数据处理结果发现了 45 条检索词条的组合，一旦将它们用于一个数学模型，它们的预测与官方数据的相关性高达 97%。

数据往往都是不完美的，拼写错误和不完整短语很普遍。为什么谷歌可以实现这么精准的预测?如果从因果关系看，是因为人感到不舒服，或听到别人打喷嚏，或者阅读了相关的新闻后感到焦虑吗?谷歌不是从这种因果关系去考虑的，而是从相关性的角度去预测一个持续发展的大方向，因为大众的搜索词条处于不断变化之中，外界的一个蝴蝶翅膀的扇动，就会使搜索发生系统的、混沌的变化。

英国华威商学院的研究人员与波士顿大学物理系的研究人员合作，同样通过谷歌趋势（Google Trends）服务预测股市的涨跌。研究人员使用谷歌趋势共计追踪了 98 个搜索关键词，其中包括“债务”、“股票”、“投资组合”、“失业”、“市场”等与投资行为相关的词，也包括“生活方式”、“艺术”、“快乐”、“战争”、“冲突”、“政治”等与投资无关的关键词，发现有些词条，如“债务”成为预测股市的主要关键词，这篇题为《使用谷歌趋势量化金融市场的交易行为》（*Quantifying Trading Behavior in Financial Markets Using Google Trends*）的论文也发表在《自然》杂志上。同样，2010 年，美国印第安纳大学的研究人员也发现：Twitter 用户的情绪有助于预测股市。诺贝尔经济学奖获得者罗伯特·席勒所倡导的“动物精神”，在大数据的相关性检验下，可以对资产定价实现预测。

当然，谷歌的算法并不是百试百灵的。例如，“谷歌流感趋势”曾经显示，有 10%的美国人可能患上了流感。但美国疾病控制和预防中心的数据显示，峰值只有 6%左右。经过研究发现，原来这是因为谷歌的算法未能充分考虑一些新的外部影响因素所致。例如，媒体对

流感的报道增多和社交媒体对流感的讨论增加，都会对该服务的数据和统计信息产生影响。流感新闻大爆炸在很大程度上改变了人们的搜索词条。这使人联想到物理学中经典的“测不准原理”。物理学家玻尔认为在量子理论中，任何对原子体系的观测都会涉及所观测对象在观测过程中的改变，和谷歌的算法一样，我们自身的行为可能也在谷歌的观测中改变，因此不可能对量子有单一的定义，也不可能用平常所谓的因果性去理解谷歌预测的趋势。

资料来源：中国大数据. http://www.thebigdata.cn/YeJieDongTai/11333.html

第9章 统计分析与统计报告

学习目标要求

1. 了解统计分析的概念和形式，掌握统计分析的步骤和方法。

2. 掌握统计分析报告的含义和作用，了解统计分析报告的结构，并能结合实际资料撰写统计分析报告。

重点难点提示

1. 统计分析的步骤，统计分析的方法。

2. 统计分析报告的结构，撰写分析报告。

教学引导案例

2016年上半年旅游统计数据报告及下半年旅游经济形势分析

一、上半年旅游统计数据报告

2016年上半年，我国旅游市场规模稳步扩大，继续领跑宏观经济：国内旅游22.36亿人次，比上年同期增长10.47%；出入境旅游1.27亿人次，增长4.1%；上半年实现旅游总收入2.25万亿元，增长12.4%。

（1）上半年国内旅游人数增长10.47%

根据国内旅游抽样调查结果，2016年上半年，国内旅游人数22.36亿人次，比上年同期增长10.47%。其中，城镇居民15.17亿人次，增长13.55%；农村居民7.19亿人次，增长4.51%。国内旅游收入1.88万亿元，增长13.72%。其中城镇居民花费1.48万亿元，增长15.13%；农村居民花费0.40亿元，增长8.84%。

（2）上半年入境旅游人数和入境过夜旅游人数分别增长3.8%和4.3%

2016年1~6月，入境旅游人数6787万人次，比上年同期增长3.8%。其中，外国人1347万人次，增长9.0%；香港同胞4 003万人次，增长2.2%；澳门同胞1 158万人次，增长3.5%；台湾同胞279万人次，增长5.8%。入境旅游人数按照入境方式分，船舶占3.4%，飞机占16.0%，火车占0.8%，汽车占21.5%，徒步占58.3%。

2016年1~6月，入境过夜旅游人数2 887万人次，增长4.3%。其中，外国人1 036万

人次，增长 6.8%；香港同胞 1 369 万人次，增长 2.3%；澳门同胞 236 万人次，增长 3.5%；台湾同胞 246 万人次，增长 6.0%。

（3）上半年国际旅游收入达 570 亿美元

2016 年 1～6 月，国际旅游收入 570 亿美元，比上年同期增长 5.3%。其中，外国人在华花费 310 亿美元，增长 6.9%；香港同胞在内地花费 151 亿美元，增长 2.3%；澳门同胞在内地花费 38 亿美元，增长 3.5%；台湾同胞在大陆花费 72 亿美元，增长 6.0%。

（4）上半年入境外国游客亚洲占比 63.7%，以观光休闲为目的游客占 31.2%

2016 年 1～6 月，入境外国游客人数 1 347 万人次，亚洲占 63.7%，美洲占 12.4%，欧洲占 18.9%，大洋洲占 2.9%，非洲占 2.0%，其他国家占 0.1%。其中，按照年龄分，14 岁以下人数占 3.8%，15～24 岁占 7.7%，25～44 岁占 45.8%，45～64 岁占 36.6%，65 岁以上占 6.1%；按性别分，男占 64.5%，女占 35.5%；按目的分，会议/商务占 20.2%，观光休闲占 31.2%，探亲访友占 3.4%，服务员工占 14.8%，其他占 30.4%。

2016 年 1～6 月，按入境旅游人数排序，我国主要客源市场前 17 位国家如下：韩国、日本、美国、越南、俄罗斯、菲律宾、蒙古、马来西亚、新加坡、印度、泰国、加拿大、澳大利亚、德国、印度尼西亚、英国、法国。

（5）上半年中国公民出境旅游人数达 5 903 万人次

1～6 月中国公民出境旅游人数累计 5 903 万人次，比上年同期增长 4.3%。

二、下半年旅游经济形势分析

（一）旅游消费趋势

1. 国内旅游继续保持两位数增速

旅游业将继续领跑宏观经济。我国经济进入结构深度调整的关键期，传统货币政策和财政政策效应弱化，经济增速将缓慢筑底，企稳后继续呈 L 形走势。在此背景下，旅游一枝独秀的地位将更加明显。

旅游需求升级空间更加凸显。目前，观光、休闲等传统旅游消费稳定发展，度假旅游受制于供给缺口、休假时长偏短、多数地区经济发展水平偏低等因素，消费升级、拉动经济和就业增长的效应将更加明显。虽然三季度开始极端天气频发，观光休闲和度假旅游需求均受到不同程度的抑制，但是下半年将会出现恢复性增长。

2. 入境旅游有望达到 4%的良好增长势头

在人民币贬值、空气质量改善、外围目的地恐怖活动频发等因素的推动下，入境旅游由前几年的衰退周期转为复苏周期的基础进一步稳固。下半年入境旅游人次增长有望突破 4%，国际旅游收入增长超过 5%。

3. 出境旅游维持 4%左右的增速

经济下行和人民币贬值背景下旅游的“进口替代”加速显现，即国民倾向于用国内游替代出境游。考虑到出境旅游已经进入国民常态化生活选项，加上研学旅行、邮轮旅游、国庆度假等积极因素，下半年出境旅游还会有 4%左右的增长。

（二）旅游投资分析

经济下行期间，大量具有准公共物品性质的旅游投资易获青睐，旅游投资继续快速增长的势头可以保持，但宏观向好的同时也要关注微观层面的利空信号。

旅游上市公司板块上半年市场表现较弱，跌幅 27.26%，跑输大盘 10%。景区、旅行社及酒店餐饮板块跌幅均超过 25%。大批旅游创新创业公司进入瓶颈期，淘在路上、麦兜旅行、

拒宅网、脚丫旅行网、找好玩、周五旅行网等创业公司宣布破产或倒闭。上半年在线旅游领域投资事件为 57 起，仅为去年全年的 30%。

（三）涉旅政策

今年以来，《国家创新驱动发展战略纲要》、《国务院办公厅关于加强旅游市场综合监管的通知》、《国务院关于深化投融资体制改革的意见》、《国家旅游局关于旅游不文明行为记录管理暂行办法》、《关于开展特色小城镇培育工作的通知》等文件陆续出台，将从不同方面对旅游发展形成综合助推力。同时也应看到，经济结构深度调整期政策空转率较高，强化落实，特别是推动政策效益直接起效在旅游实体经济中是关键。

（四）旅游就业测算

预计，2016 年我国旅游直接就业人口总数约为 2 820 万，直接和间接旅游就业人口总和约为 7 974 万人，对全国就业的综合贡献率为 10.25%。

资料来源：http://travel.sohu.com/20160801/n462096235.shtml

思考：

1. 通过分析报告的数据，你对中国旅游业发展有什么看法？为什么？

2. 中国旅游业快速增长的原因是什么？

第一节　统计分析概述

一、统计分析的概念、作用和特点

（一）统计分析的概念

统计分析，就是运用各种统计综合指标和方法，将丰富的统计资料和生动的具体情况结合起来，对社会经济现象的各个方面进行分析研究，从而揭示其发展变化的规律性，提出解决问题的方法的一种逻辑思维活动。

（二）统计分析的作用

统计分析是整个统计工作的一个重要阶段，是统计工作的最终环节，是充分发挥统计整体职能的关键环节，其好坏直接影响统计的质量。在统计实践中，只有开展统计综合分析，才能更好地发挥统计的作用，为各级领导和有关方面的公众提供有数据、有分析的资料，为制订计划和规划、实行宏观调控、决定有关方针和政策提供科学依据。具体表现在以下几个方面。

1. 全面、准确地反映客观情况

统计分析从数量上、总体上认识客观事物，既是人们认识上的需要，使认识更加清晰、明确，又可以避免以偏概全，使人们的认识较为全面和正确。

2. 深入把握社会经济现象的规律

只有对客观现象总体的数量方面进行分析，才能获得规律性的认识。

3. 参与社会经济管理

统计分析把数据、情况、问题、建议等融为一体，既有定量分析又有定性分析，比一般统计数据更集中、更系统、更清楚地反映客观实际，便于研究、理解和利用，因而是发挥统计信息、咨询、监督作用的主要手段。与此同时，也提高了统计工作的社会地位。

（三）统计分析的特点

1. 以统计数据为基础，定量与定性分析相结合

统计数字是统计分析最主要的源泉，是统计分析的主要语言。一篇好的统计分析报告通常自始至终利用统计数字进行分析。

2. 综合运用多种分析方法

统计分析要采用分组分析法、对比分析法、集中趋势分析法、离中趋势分析法、指数因素分析法、相关分析法、综合评价分析法等多种统计方法进行分析。

3. 统计分析要将数字与实际情况相结合

统计分析不能离开实际情况单纯地罗列数字。有些情况不可能用统计数字来反映，若单凭统计数字进行分析，往往难于把问题说清楚，也不易把问题产生的原因和情况弄明白，因此要密切结合数字背后的实际情况进行分析，以便得出具体而确切的分析结论。

4. 统计分析要具有时效性

这是保证统计信息价值的重要条件，不适时的统计分析是无意义的活动。

二、统计分析的形式

（一）按照研究的内容划分，有综合分析和专题分析

1. 综合分析

综合分析是从整体上对研究对象的各个方面带有全局性的问题进行的系统性分析研究，如国民经济综合分析、部门综合分析、地区综合分析、企业综合分析等。

2. 专题分析

专题分析是对社会经济发展的某一方面、某一环节的现象中重大问题或关键问题所进行的专门分析和研究，如某企业的销售分析、某地区可持续发展分析等。

（二）按照研究对象的层次划分，有宏观、中观和微观统计分析

1. 宏观分析

宏观分析主要是指对国民经济的发展目标和总任务、战略重点、战略步骤、总量变动及发展规律等问题所进行的分析。国民经济要保持长期、稳定、协调、持续的发展，就必须长远规划。

2. 中观分析

中观分析指针对地区、部门的长远发展规划和目标，通过对大量、全面的资料的研究和统计分析，可以为制定区域和行业的发展战略和制定有关方针政策提供依据，促进各项事业的发展。

3. 微观分析

微观分析主要是指对个别消费者、生产者、企业和单位的经济活动的数量变化和发展规律所进行的统计分析。在实际工作中，运用最广泛的是对企业经营管理、产品经营和经济效益方面的分析研究。

（三）按照观察时间不同划分，有定期分析和预计分析

1. 定期分析

定期分析是指在一定时期内对生产经营活动情况的全面分析。

2. 预计分析

预计分析是指在报告期尚未结束前，根据计划完成进度，结合主观条件，预计到计划期结束前任务完成情况的一种分析。

此外，按照分析的类型划分，还可以划分为调查型、说明型、情报型、公报型、研究型、预测型和报道型分析等；按照统计认识作用的层次不同，可分为状态分析、规律分析和前景分析。统计分析的不同形式反映的统计分析的内容和要求都不一样，因此应根据不同情况搞好分析研究，写好分析报告，体现统计分析的广泛性和多样性特征，防止统计分析的公式化和固定化。

三、统计分析的一般步骤

统计分析的一般步骤如下。

（一）选择并确定研究课题

选择课题是指从客观存在的现实和大量的统计资料中选择所要研究和反映的对象，确定研究目的和范围，规划主题思想和基本内容。准确地选择研究课题，是统计分析成功的关键。选择准确可以提高统计分析的价值，并为以后的取材、构思、表达等打下一个良好的基础。

从实际工作来看，选题可以围绕以下几方面：一是要抓住领导关心的问题，特别是领导亲自出的选题；二是要抓具有现实意义的和与全局性工作有密切联系的课题；三是要抓经济社会发展中带有苗头性、突发性的问题；四是要抓改革开放和经济建设中的新情况、新问题、新经验；五是要抓各方面有不同看法的重大问题；六是要配合中心工作、重要会议提供的材料。

（二）拟定分析提纲

分析提纲要紧扣分析目的和分析题目，分析提纲一般包括下列内容：

①统计分析的目的和要求；

②从哪些方面进行分析，要列出分析大纲及分析的细目；

③从哪些方面收集资料，以及资料的来源；

④收集资料的方式和方法。

（三）收集鉴别与整理资料

统计分析的依据是统计资料，选择客观、详尽、代表性强、有说服力的统计资料是搞好统计分析的又一重要前提。对于这些资料，需要分析人员注意通过收集抄录、现场调查、日常观察、广泛阅读等来获取。

统计分析所使用的资料内容广泛，来源渠道多，包括：

①统计资料；

②调查情况，指在统计调查中取得的活动情况，有文字记载，也有没有文字记载需要统计分析人员去记录的；

③见闻资料；

④政策法规和有关名人言论等。

在整理资料时，要注意收集的资料是否符合客观实际，是否新颖，是否具有普遍意义。而且要注意数字资料的时效性、可比性、完整性等。对于由于种种条件的限制，不能直接取得而分析又必须使用的数字资料，可用统计中的各种估算方法来取得。

（四）运用各种方法进行系统周密的分析

运用各种统计分析方法进行系统而严密的分析研究，如综合指标法、分组法、因素分析法、动态分析法等方法对资料进行加工整理、归纳分析，从定量入手，以求达到对质的认识。

（五）得出结论，提出建议

通过对事物的现象和解析，并进行各种各样的比较分析后，应对所分析的问题做出实事求是的结论，结合有关的数据、综合比较的结果，从中揭示其发展演变的规律性，对存在的问题提出建议。

（六）根据分析结果形成分析报告

统计分析报告是在对统计资料和有关情况进行研究的基础上，用简洁明确的文字叙述研究过程、表述分析研究结果的一种主要形式。

第二节　统计分析方法

一、统计分析方法概述

前面有关章节介绍的各种分析方法，如分组法、静态指标法、动态分析法、指数法、相关法和方差分析法等属于传统的统计分析方法。科学地在统计分析工作中运用这些方法，对现象进行综合分析研究，能够全面、深入地认识问题。现从分析综述角度对常用分析方法进行概要说明。

（一）对比分析法

对比分析法是把客观事物加以比较，以达到认识事物的本质和规律并做出正确的评价。对比分析法通常把两个相互联系的指标数据进行比较，从数量上展示和说明研究对象规模的大小、水平的高低、速度的快慢及各种关系是否协调。在对比分析中，选择合适的对比标准是十分关键的步骤，选择合适，才能做出客观的评价，选择不合适，评价可能得出错误的结论。对比标准存在以下几种选择。

1. 时间标准

时间标准即选择不同时间的指标数值作为对比标准，最常用的是与上年同期比较（同比），还可以与前一时期比较，此外还可以与达到历史最好水平的时期或历史上一些关键时期进行比较。

2. 空间标准

空间标准即选择不同空间指标数据进行比较。

①与相似的空间比较，如本市与某些条件相似的城市比较。

②与先进空间比较，如我国与发达国家比较。

③与扩大的空间标准比较，如我市水平与全国平均水平比较。

3. 经验或理论标准

经验标准是通过对大量历史资料的归纳总结而得到的标准，如衡量生活质量的恩格尔系数、反映社会收入差距程度的基尼系数等。理论标准则是通过已知理论经过推理得到的依据。

4. 计划标准

计划标准即与计划数、定额数、目标数对比，市场经济并不排斥科学合理的计划，因此计划标准对统计评价仍有一定的意义。

相联系的两个指标对比，表明现象的强度、密度、普遍程度，如人均国内生产总值、人口密度、人均收入及某些技术经济指标等。对比分析按说明的对象不同可分为单指标对比（简

单评价）、多指标对比（综合评价）。在进行对比分析时应掌握下面几个原则：指标的内涵和外延可比；指标的时间范围可比；指标的计算方法可比；总体性质可比。

（二）结构分析法

结构分析法是在统计分组的基础上，计算各组成部分所占比重，进而分析某一总体现象的内部结构特征、总体的性质、总体内部结构依时间推移而表现出的变化规律性的统计方法。结构分析法的基本表现形式就是计算结构指标，这是指总体各个部分占总体的比重，因此总体中各个部分的结构相对数之和等于 100%。

通过结构分析可以认识总体构成的特征。例如，某地区近五年来高新技术产品比重第一年占 20%，第三年占 32%，第五年占 51%，表明产业结构向高新技术产业的转变。结构分析也可以揭示现象之间的依存关系，如研究商业企业中商品销售额与流通费用的依存关系，可将各商品销售额分组，计算每个组相应的商品流通费用。例如，某市年销售额 300 万元以上的企业占 15%，每万元商品销售额中的流通费为 6.0 元，而 300 万元以下的企业占 85%，流通费为 8.5～11.2 元，说明销售规模越大的企业流通费用越少。



案例学习 9-1

在世界 500 强企业最多的前五个国家中，其 500 强企业的数量、营业额、利润等经济指标对比情况如表 9-1 所示。

表 9-1　世界 500 强企业最多的前五个国家的情况

位次(按企业数排列)		国家	企业数（家）		营业额			利润额		
					总计（亿美元）	占世界 500 强企业的比重（%）		总计（亿美元）	占世界 500 强企业的比重（%）	
2015 年	2014 年		2015 年	2014 年		2015 年	2014 年		2015 年	2014 年
1	1	美国	162	170	73 383	35.1	36.0	5 881	38.5	37.5
2	2	日本	67	70	24 072	11.5	12.3	1 066	7.0	8.8
3	4	法国	38	38	18 108	8.7	8.2	1 125	7.4	7.0
4	5	德国	37	35	18 365	8.8	8.5	864	5.3	8.0
5	3	英国	33	38	15 448	7.4	8.7	1 264	8.3	5.5

分析：本案例利用结构分析法说明世界 500 强企业最多的前五个国家的企业数、营业额、利润额等方面的情况。

（三）平均和变异分析法

平均和变异分析法是利用平均指标和变异指标分析社会经济现象的一般水平及差异的方法。

平均指标是同质总体中各单位某一指标值的平均数字，反映总体在一定时间、地点条件下的一般水平，如平均工资、单位产品成本、单位面积产量、平均单价等。变异指标则是说明总体各单位标志值差异程度的指标，常用的变异指标是标准差和标准差系数。

使用平均和变异分析法应注意以下几点：①正确的计算平均指标必须是同质总体的平均数；②平均指标与变异指标结合运用，全面认识和评价总体，既能说明总体的一般水平，又能说明总体内部差异的程度，如甲单位月平均工资 1 600 元，标准差为 60 元，标准差系数为 3.75%（即 60÷1 600×100%），乙单位月平均工资 800 元，标准差为 40 元，标准系数为 5%（即 40÷800×100%），说明甲单位工资水平高于乙单位，差异程度低于乙单位，平均工资的代表性高于乙单位；③用组平均数补充总平均数，正确认识总体结构对平均水平的影响；④结合典型事例进行分析。

（四）动态分析法

动态分析法是以客观现象所显现出来的数量特征为标准，判断被研究现象是否符合正常发展趋势的要求，探求其偏离正常发展趋势的原因并对未来的发展趋势进行预测的一种统计分析方法。

动态分析法主要包括两个方面：

第一，编制时间数列，观察客观现象发展变化的过程、趋势及其规律，计算相应的动态指标，用以描述现象发展变化的特征；

第二，编制较长时期的时间数列，在对现象变动规律性判断的基础上，测定其长期趋势、季节变动的规律，并据此进行统计预测，为决策提供依据。

观察编制好的时间数列，可以看出现象变化的大致过程和趋势，但要给予定量分析，必须计算各种动态分析指标：一类是动态比较指标，主要有增长量、发展速度、增长速度；一类是动态平均指标，主要有平均发展水平、平均发展速度、平均增长速度。时间数列的形成是各种不同影响事物发展变化的因素共同作用的结果，为了便于分析事物发展变化规律，通常将时间数列形成因素归纳为以下四类：长期趋势、季节变动、循环波动、不规则变动。

（五）平衡分析法

所谓平衡就是各个相互联系的因素之间在数量上保持一定的合理的对应关系。平衡分析法是分析事物之间相互关系的一种方法。它分析事物之间发展是否平衡，揭示事物间出现的不平衡状态、性质和原因，指引人们去研究积极平衡方法促进事物的发展。统计平衡分析的主要方法有编制平衡表和建立平衡关系式。

平衡表与一般统计表的区别在于：指标体系必须包括收入与支出、来源与使用两个对应平衡的指标。平衡表的主要形式有三种，即收付式平衡表、并列式平衡表和棋盘式平衡表，前两种形式如资产负债表、能源平衡表，后一种形式如投入产出表。

平衡关系式是用等式表示各相关指标间平衡关系的式子。例如，期初库存+本期入库=本期出库+期末库存，资产=负债+所有者权益，增加值=总产出-中间投入。

统计中的平衡分析基本要求和特点如下：

①通过有联系指标数值的对等关系来表现经济现象之间的联系；

②通过有联系指标数值的比例关系来表现经济现象之间的联系；

③通过任务的完成与时间进度之间的正比关系来表现经济现象的发展速度；

④通过各有关指标的联系表现出全局平衡与局部平衡之间的联系。

（六）相关分析法

相关分析法是测定经济现象之间相关关系的规律性，并据以进行预测和控制的分析方法。实践中进行相关分析主要依次解决以下问题。

①确定现象之间有无相关关系及相关关系的类型。

对不熟悉的现象，则需收集变量之间大量的对应资料，用绘制相关图的方法进行初步判断。从变量之间相互关系的方向看，变量之间有时存在着同增同减的同方向变动，是正相关关系；有时变量之间存在着一增一减的反方向变动，是负相关关系。从变量之间相关的表现形式看有直线相关和曲线相关，从相关关系涉及变量个数看，有一元相关（或简单相关关系）和多元相关（或复相关关系）。

②判定现象之间相关关系的密切程度。

通常计算相关系数 R，必要时应对 R 进行显著性检验，相关系数绝对值在 0.8 以上表明高度相关。

③拟合回归方程。

如果现象间相关关系密切，就根据其关系的类型建立数学模型，用相应的数学表达式——回归方程来反映这种数量关系，这就是回归分析。

④判断回归分析的可靠性。

要用数理统计的方法对回归方程进行检验。只有通过检验的回归方程才能用于预测和控制。

⑤根据回归方程进行内插外推预测和控制。

（七）综合评价分析法

随着统计分析活动的广泛开展，评价对象越来越复杂，简单评价方法的局限性也越来越明显。经常会出现从这几个指标看甲单位优于乙单位，从那几个指标看乙单位优于丙单位，从其他指标看丙单位又优于甲单位的情况，使分析者难以评价谁优谁劣，因此通过对实践活动的总结，逐步形成了一系列运用多个指标对多个参评单位进行评价的方法，称为多变量综合评价方法，或简称综合评价方法。其基本思想是将多个指标转化为一个能够反映综合情况的指标来进行评价。例如不同国家经济实力、不同地区社会发展水平、企业经济效益评价等，都可以应用这种方法。

综合评价法的特点表现为：评价过程不是按逐个指标顺次完成的，而是通过一些特殊方法将多个指标的评价同时完成；在综合评价过程中，一般要根据指标的重要性进行加权处理，评价结果不再是具有具体含义的统计指标，而是以指数或分值表示参评单位“综合状况”的排序。

综合评价方法的步骤如下：

①确定综合评价指标体系，这是综合评价的基础和依据；

②收集数据，并对不同计量单位的指标数据进行同度量处理；

③确定指标体系中各指标的权数，以保证评价的科学性；

④对经过处理后的指标进行汇总，计算出综合评价指数或综合评价分值；

⑤根据评价指数或分值对参评单位进行排序，并由此得出结论。

综合评价分析指标值的计算方法很多，主要有打分综合法、打分排队法、综合指数法、功效系数法等。

（八）因素分析法

因素分析法是用来测定受多种因素影响的某种经济现象总变动中各个因素的影响方向和影响程度的一种统计分析方法。常用的因素分析方法主要有以下几种。

1. 相关联因素的剖析

其特点不是借助数学模型，而是根据相关因素的性质，表明其数量变化对所研究对象变动的影响关系与制约关系，从本质上讲属于经验方法。



案例学习 9-2

2016 年上半年天津市消费品零售市场活跃，社会消费品零售总额同比增长 12.7%。促进零售市场发展的有利因素是：城市居民人均可支配收入同比增长 9.1%；大型市场发展加快，亿元以上商品商场的市场交易额同比增长 7.6%；商品房销售和装饰商品热销同比分别增长 16.6%和 86.1%；春节、“五一”两个“黄金周”是零售市场增长的亮点，2 月、5 月两个月社会消费品零售额大幅度增长；私营企业、个体经济零售额同比增长 24.2%，对全市零售总额增长的贡献率达 53.4%；餐饮业零售额同比增长 15.1%。

分析：上述因素分析未用数字表达式，但确实揭示了各因素变动对现象总变动的影响关系，可以借以认识事物的性质，为解决实际问题、推动事物发展提供了决策思路。

2. 指数体系及其因素分析

在经济上有联系、在数字上存在等式关系的三个或三个以上的指数称为指数体系。利用指数体系测定各影响因素对某种经济现象总体变动的方向和程度所产生的影响就是因素分析，主要有以下几种类型：对总量指标变动进行二因素分析、对总量指标变动进行多因素分析、平均指标的因素分析、指数因素分析法的扩展运用。

3. 相加因素的分析方法

在社会经济现象中，有一些现象的变动是总体内各个组成部分（或称构成因素）变动影响的结果。例如工业总产值的变动是由轻工业与重工业共同变动影响的结果。由此可见，相加因素是指现象变动是各个组成因素变动的总和，常采用比重法和差额法测定总体各个组成部分的变动。

（九）景气分析法

景气是对经济发展状况的一种综合性描述，用以说明经济活跃的程度。所谓经济景气，

是指总体经济呈上升发展趋势，呈现市场繁荣、经济总量增长速度加快的景气状态。经济不景气是指总体经济呈下滑的发展趋势，绝大部分经济活动处于收缩或半收缩状态，出现市场疲软、经济增长速度停止或迟缓、许多企业破产或倒闭、失业人数增加等现象。

景气分析法是统计分析中的一项重要内容，其研究对象是市场经济条件下宏观经济的波动，以及经济运行过程中交替出现的扩张和收缩、繁荣与萧条、高涨与衰退现象。这种分析可以帮助宏观决策部门把握国民经济运行态势；帮助生产经营部门判断宏观经济走势；还可以帮助社会公众监测宏观和微观经济运行状况。

景气分析方法主要采用时间序列分析方法、调查分析法、经济计量学方法和直

接度量法等分析方法，还采用灵活多样的景气调查方法作为总量分析方法的补充。

第三节　统计分析报告

统计分析报告就是根据统计学的基本原理和方法，运用大量统计数据来研究和反映社会经济活动的状况、成因、规律和结论的一种“报告”。从写作角度看，“报告”是一种文体，它是对某种特定的客观现象，经过某种形式研究后得出的结论性认识的文字表达；其次，统计分析报告是一种“分析”报告，这种文体，将把某种特定的客观现象的分析过程和结论表达出来：另外，统计分析报告是“统计”分析报告，“统计”一词实际上反映了统计分析报告的“质”的规定性。

一、统计分析报告的特点

1. 运用一整套统计特有的科学分析方法

运用的方法如对比分析法、动态分析法、因素分析法、统计推断等，同时结合统计指标体系，全面、深刻地研究和分析社会经济现象的发展变化。

2. 运用数字语言

运用数字语言（包括运用统计表和统计图）来描述和分析社会经济现象的发展情况，让统计数字说话，通过确凿、翔实的数字和简练、生动的文字进行说明和分析。

3. 注重定量分析

利用统计部门的优势，从数量方面来表现事物的规模、水平、构成、速度、质量、效益等情况，并把定量分析与定性分析结合起来。

4. 具有很强的针对性

针对各级党政领导和社会各界普遍关心的难点、热点、焦点问题进行分析，只有这样才能有的放矢，针对性强。

5. 注重准确性和实效性

准确性是统计分析报告乃至整个统计工作的生命，统计分析报告的准确性除了数字准确（不能有丝毫差错）、情况真实（不能有虚假）之外，还要求论述有理（不能违反逻辑）、观点正确（不能出现谬误）、建议可行（不能脱离实际）。

统计分析报告具有很强的实效性。失去了时效性，也就失去了实用性，统计分析报告写得再好，也成了无效劳动。要保证统计分析报告的时效性，统计人员要有“一叶知秋”、“见微知著”的敏感性，要有争分夺秒的时间观念。争取“雪中送炭”，避免“雨后送伞”，把统计分析报告提供在领导决策之前和社会各界需要之时。

6. 具有很强的实用性

统计分析报告是统计工作的最终结果，它不但包含了统计数据反映的信息，更为重要的是，它还能进行分析研究，能进行预测，能指出工作中的不足和问题，能提出有利于今后工作的措施和建议，从而直接满足党政领导和社会各界在了解形势、制定政策、编制计划、经营管理、检察监督、总结评论、科研教学等方面的实际需要。

二、统计分析报告的作用

1. 衡量统计工作水平的综合标准

一般来说，高质量的统计分析报告来自高质量的统计设计、统计调查、统计整理、统计分析和统计分析写作。但是，如果仅有较好的水平，统计设计、统计调查、统计整理、统计分析都是低质量的，也不可能产生高质量的统计分析报告，因此，统计分析报告写不好，当然是统计工作水平不高的表现。

2. 传播统计信息的有效工具

现代社会是信息的时代，信息已成为重要资源。统计信息又是社会信息的主体，而且是最全面、最稳定、较准确的信息。统计信息要通过载体传播，而统计分析报告是主要载体之一，适合于报纸杂志上发表，传播条件比较简便，具有较大的信息覆盖面，是传播统计信息的有效工具。

3. 决策的重要依据

现代社会经济管理必须科学决策，而科学的决策又必须依据准确、真实的统计数据。统计分析报告把原始资料信息加工成决策信息，它比一般的统计资料更能深入地反映客观实际，更便于党政领导和社会各界接受和利用，因而统计分析报告是党政领导决策的重要依据。

4. 统计服务与统计监督的主要手段

统计分析报告把数据、情况、问题、建议等融为一体，既有定量分析又有定性分析，比一般的统计数据更集中、更系统、更鲜明、更生动地反映了客观实际，又便于人们阅读、理

解和利用，是表现统计成果的好形式，是传播统计信息的有效工具，自然也就成了统计服务与统计监督的主要手段。

三、统计分析报告的分类

统计分析报告的应用是很广泛的，由于它是报告社会经济情况的一种文体，因而属于应用文范畴。统计分析报告可以从不同角度划分成很多种类。

①按统计领域划分，可分为工业、农业、商业、科技、教育、文化、卫生、体育、人口、财政、金融、政法、人民生活、国民经济综合、核算等统计分析报告。

②按写作对象的层次划分，可分为微观、中观和宏观统计分析报告。对于微观、中观、宏观的划分，目前尚无统一的标准。一般来讲，基层企事业单位、村、家庭及个人，属于社会经济的“细胞”，可视为“微观”；乡镇、县一级可视为“中观”；而地（市）及地（市）以上的地区和部门，由于地域较广，社会经济门类比较复杂，需要较多地注意平衡关系，可视为“宏观”。

③按内容范围分，可分为综合与专题统计分析报告。

综合统计分析报告是研究和反映一个地区、部门或单位全面情况的分析报告，这种分析报告一般是定期的。所谓综合，既包括各方面的意思，也包含着综合方法的意思。

专题统计分析报告是研究和反映某一方面或某个专门问题的分析报告。专题统计分析报告有定期的，也有不定期的，以不定期的较多。

④按照时间长度分，可分为定期与不定期的统计分析报告。

定期统计分析报告一般利用当年的定期统计表制度的统计资料来定期研究和反映社会经济情况。根据期限不同，定期统计分析报告又可分为日、周、旬、半月、月度、季度、上半年、年度等统计分析报告。

不定期的统计分析报告主要用于研究和反映不需要经常性定期调查的社会经济情况。

⑤按写作类型分，可分为说明型、快报型、计划型、总结型、公报型、调查型、分析型、研究型、预测型、资料型、信息型、微型、综合型、文学型、系列型等十五种类型的统计分析报告。



知识拓展

统计分析报告与统计调查报告的区别

统计研究的成果既可以用统计分析报告的形式表述，也可以用统计调查报告的形式表述。在统计工作中，这是使用最多的两种专业应用文。它们在发挥统计工作的作用上、在文

体的写作要求上有许多相同的地方，都以社会经济现象作为研究和反映的对象，都要运用统计调查得来的数字资料分析事物的数量关系，都为领导提供决策咨询，在写作程序和写作方法上也基本相同。于是就产生了统计分析报告与统计调查报告不分之说，而且在实际工作中确也存在着一些不能明确区分到底是哪种文体的情况。但是，我们只要认真地研究，仔细地比较，就可以发现，它们不管在材料的来源和使用上，还是在表现的形式上，都各有侧重，从而也就体现出了它们的特点和区别。

首先，从材料的来源和使用上看，虽然运用各种统计调查方法所收集、整理的统计资料都是这两种报告材料的来源，但是统计分析报告主要依靠全面统计报表和抽样调查所提供的数字资料来进行分析，以间接材料为主，辅之以实际调查的情况。而统计调查报告则主要是从各种专门调查、实际调查中去获取资料，要求作者必须亲自深入实际进行细致调查，因此它以直接材料为主。在使用材料时，统计分析报告要求充分利用全面统计数字，就是抽样调查也要推算总体情况，而典型材料的运用较少，有时甚至没有。而统计调查报告要以典型材料为主，有时甚至以一个典型事件贯穿全文的始终。通过对有代表性的典型事实的深入解剖来认识事物发展变化的过程和趋势。一般只反映部分单位的情况，不直接反映或推算总体情况。因此，统计调查报告在材料使用上比统计分析报告更细致、具体。

另外，从内容的侧重和表现形式上看，统计分析报告一般不交代调查活动的基本情况，更不反映调查的具体过程，而是侧重于分析，运用统计分析的各种方法，从数字分析入手，揭示事物的数量关系，分析现象产生的原因，提出解决问题的办法，因此它的说理性和说明性较强。而统计调查报告侧重于反映调查情况，介绍调查对象，一般要先交代调查的目的、对象、内容及调查的简要过程，然后运用大量的事实反映调查的结果，最后得出结论，探讨规律，因此它的陈述性和报导性较强。此外，统计分析报告的写法灵活，可长可短，形式多样，而统计调查报告的写法则相对比较规范和完整。

四、统计分析报告的结构

统计分析报告大致包括：标题、导语、主体、结尾四部分。

（一）标题

恰当地选定标题，是进行统计分析的重要一步。常见的标题拟定形式如下。

1. 论点题

论点题能揭示主题，摆明观点，如《我省在全国经济发展中的战略地位和作用》、《调整产业结构是农村富裕的必由之路》等。

2. 设问题

设问题能引起读者疑问、思考，刺激读者的阅读欲望，如《住房为什么紧张？》等。

3. 比喻、对比、加重语气题

这种题目能通过对比引人注意，如《“骨之不强”，肉将焉附？——谈投资结构问题》，显得新颖别致，醒目强烈。

（二）导语

导语是统计分析报告内容的引语，是整个分析报告的开头。它是关系到分析报告成效的一个重要因素，因此，对导语的基本要求，一是要能够吸引读者，使读者有读下去的兴趣；二是要为全文的展开理清脉络，牵出头绪，确定格局；三是要短、精、新。统计综合分析报告中常用的导语形式如下。

1. 开门见山，揭示主题

其特点是简明扼要，直叙入题，这种导语是统计分析报告最常用的形式之一。例如，“根据全国人大财经委的通知要求，现对当前全国经济运行的主要特点、存在问题及发展趋势作一简要分析，供参考。”

2. 纵览全文的观点

例如，“4 月份，江苏省货币信贷运行特征变动明显，各项贷款平稳增长，储蓄存款大幅下降，企业存款定期化增强；各项贷款增势明显回落……流动性水平略降。”

3. 交代写作的目的或动机

这也是目前常用的开头方式之一。这种开头的主要特点是：起因线索完整，时间、地点俱在，分析动机清楚，命题明显自然。例如，“充分研究这一时期我市工业性投资的发展状况及存在的问题，将对我市在实现 GDP 千亿元规划目标的过程中，如何让合理确定工业性投资的规模、结构及与其他投资的比例关系提供有益的借鉴。为此，我们对“七五”以来我市工业性固定资产投资情况进行了如下分析。”

4. 突出矛盾，造成悬念

这是旨在分析问题或阐述观点之前，提出一个问题，以引起读者的注意和思考。例如，“……然而，从目前的情况来看，达到全国平均水平的任务较为艰巨。那么，究竟需要多长时间才能实现这一目标，两个五年计划？还是更长一些时间？对此，我们在全面分析某省与全国人均 GDP 差距的基础上，就达到全国平均水平所需的时间和速度进行了初步测算，并对某省经济实现赶超的发展思路进行梳理，以期为省领导和有关经济管理部门指导经济工作提供参考。”

5. 设靶论战

即有意识地设置“对立面”，把不同观点列举出来，然后加以评论。

例如，“这几年来，国民经济增长速度不断趋缓，引起国内外社会各界对我国未来经济发展趋势的极大关注。我国经济究竟是由快速发展进入低速发展呢？还是继续保持快速发展？双方各持己见，分歧很大，对此，本文就这一问题作一初步分析。”

此外，统计综合分析报告也可不写导语，直接进入主体部分。

（三）主体

主体结构的形式具体体现在层次段落上，层次即指内容的先后次序，常见的有如下几类。

1. 叙时连贯式

按事物发展经过和时间顺序安排层次，各层意思之间是连贯的关系。

2. 叙事递进式

文章各部分内容按事理的发展顺序排列。它可以是先因后果、或先果后因的因果叙事式；也可以是按事理发展的连续性，每一阶段一个层次；也可以是按事理意义上层层深入的递进关系的递进式。

3. 总分式

先总起来说，然后分开说；或者先分开说，后总起来说；或者前后都有总说，中间分开说。因分述内容不同，可以是平行总分式、对比总分式、递进总分式和叙事总分式。

4. 平列式

各部分内容相对独立，各层意思之间是平行并列关系。这种结构形式可以是同事平列式，也可以是异事平列式。

5. 简要式

一般是篇幅短小、层次简单的分析报告，多用于快报、信息、简讯。

（四）结尾

结束语是统计分析报告的结尾，它是文章思想内容的必然归宿，一个好的文章结尾可以帮助读者明确题旨、加深认识，引起读者的联想和思考，对结尾的要求是自然、简短有力。统计分析报告的结尾写法没有硬性规定，主要由文章内容决定，要不落俗套，不断创新。

综合统计分析报告结尾常见的写法如下。

1. 总括全文，呼应开头

报告在论证观点、结束全文之时予以归纳总结，突出中心思想，呼应主题。例如，“通过对以上经济效益指标的定量分析，说明工业品出厂价格和燃料购进价格的变动已成为影响工业经济效益的主要因素。因此，逐步理顺工业品的价格关系，已成为当前一个十分重要的问题。”

2. 强调看法和建议

以建议结束全文也是统计综合分析报告常见的方式，或没有结尾段，以最后一层次的若干建议来收笔，或专门有一个建议结尾段，用总结建议内容的方式收尾。例如，“既然在 10.2% 的速度下经济增长的产业协调性还有所增强，经济增长的动力仍具有可持续性，经济发展的瓶颈约束有所减缓，经济增长的效益有明显提高，经济增长的金融环境整体平稳，那就不要轻言 10.2%的速度“过热”了，更不要草率做出宏观调控政策需要紧缩的结论。”

3. 对未来进行展望

以积极的心态提出新问题，展示发展前景，预测未来发展趋势。例如，“我们认为，只

要趋利避害，操作得当，2008 年我国国民经济仍可望继续保持‘高增长、低通胀’的良好运行格局。”



知识拓展

统计分析报告的修改

所谓修改，就是文章的初稿（草稿）写成后，作者对文章的内容和形式在进行多方面的加工，使其不断完善、提高，直至定稿的过程。

修改统计分析报告时需要注意以下几个问题。

①修改需要冷处理。写作统计分析报告时要争取“一气呵成”，以求文意贯通，语气畅达。但在修改的时候，却不能“趁热打铁”，而要冷处理，如果时间允许，可放一段时间，使头脑冷静后再修改，就能发现许多不妥之处。

②要正确地听取意见。“当事者迷，旁观者清”，一方面听取领导的意见，一方面要听取其他统计人员的意见。

③抓住重点修改。修改当然要推敲某些字句，但这不是重点，修改的重点主要是核实数字和情况是否准确、观点是否正确、论据是否充分、说理是否透彻、意义是否深刻。



本章小结

统计分析是运用各种统计综合指标和方法，将丰富的统计资料和生动的具体情况结合起来，对社会经济现象的各个方面进行分析研究，从而揭示其发展变化的规律性并提出解决问题的方法。统计分析是整个统计工作的最终环节，是充分发挥统计整体职能的关键阶段，直接影响统计工作的质量。

在统计分析工作中应科学地运用统计分析方法对现象进行综合分析研究，能够全面、深入地认识问题。常用的分析方法有对比分析法、结构分析法、平均和变异分析法、动态分析法、平衡分析法、相关分析法、综合评价分析法、因素分析法和景气分析法。

统计分析报告是根据统计学的基本原理和方法，运用大量统计数据来研究和反映社会经济活动的状况、成因、规律和结论的一种文章，其结构大致包括标题、导语、文体和结尾四部分。统计分析报告的应用十分广泛。

复习思考题

1．统计分析的概念、特点和作用是什么？
2．在统计实践中应该按照哪些步骤进行统计分析？
3．统计分析报告从不同角度可以分为哪些类型？
4．统计分析报告的写作步骤有哪些？
5．统计分析有哪些方式方法？
6．统计分析报告的概念、特点和作用是什么？

实训基础训练

1．请登录国家统计局网站或本省、市统计局网站，进入统计分析栏目，阅读统计分析报告，并结合本章有关理论进行分析。
2．请针对本班同学的手机使用情况进行调查，搜集有关资料并撰写一份统计分析报告。
3．请结合本地区物流市场的运行状况进行调查，搜集资料，撰写统计分析报告。

课外延伸阅读

2015 年国民经济运行稳中有进、稳中有好

（2016 年 1 月 19 日）

中华人民共和国国家统计局

2015 年，面对错综复杂的国际形势和不断加大的经济下行压力，党中央、国务院保持战略定力，统筹谋划国际国内两个大局，坚持稳中求进工作总基调，主动适应引领新常态，以新理念指导新实践，以新战略谋求新发展，不断创新宏观调控，深入推进结构性改革，扎实推动“大众创业、万众创新”，经济保持了总体平稳、稳中有进、稳中有好的发展态势。

初步核算，全年国内生产总值 676 708 亿元，按可比价格计算，比上年增长 6.9%。分季度看，一季度同比增长 7.0%，二季度增长 7.0%，三季度增长 6.9%，四季度增长 6.8%。分产业看，第一产业增加值 60 863 亿元，比上年增长 3.9%；第二产业增加值 274 278 亿元，增长 6.0%；第三产业增加值 341 567 亿元，增长 8.3%。从环比看，四季度国内生产总值增长 1.6%。

一、农业生产再获丰收

全年全国粮食总产量 62 143 万吨，比上年增加 1 441 万吨，增长 2.4%。其中，夏粮产量 14 112 万吨，增长 3.3%；早稻产量 3369 万吨，下降 0.9%；秋粮产量 44 662 万吨，增长 2.3%。谷物产量 57 225 万吨，比上年增长 2.7%。棉花产量 561 万吨，比上年下降 9.3%。全年猪牛羊禽肉产量 8 454 万吨，比上年下降 1.0%，其中猪肉产量 5 487 万吨，下降 3.3%。禽蛋产量 2 999 万吨，比上年增长 3.6%；牛奶产量 3 755 万吨，增长 0.8%。

二、工业升级态势明显

全年全国规模以上工业增加值按可比价格计算比上年增长 6.1%。分经济类型看，国有控股企业增加值比上年增长 1.4%，集体企业增长 1.2%，股份制企业增长 7.3%，外商及港澳台商投资企业增长 3.7%。分三大门类看，采矿业增加值比上年增长 2.7%，制造业增长 7.0%，电力、热力、燃气及水生产和供应业增长 1.4%。新产业增长较快，全年高技术产业增加值比上年增长 10.2%，比规模以上工业快 4.1%，占规模以上工业比重为 11.8%，比上年提高 1.2%。其中，航空、航天器及设备制造业增长 26.2%，电子及通信设备制造业增长 12.7%，信息化学品制造业增长 10.6%，医药制造业增长 9.9%。全年规模以上工业企业产销率达到 97.6%。规模以上工业企业实现出口交货值 118 582 亿元，比上年下降 1.8%。12 月份，规模以上工业增加值同比增长 5.9%，环比增长 0.41%。

1～11 月份，全国规模以上工业企业实现利润总额 55 387 亿元，同比下降 1.9%。规模以上工业企业每百元主营业务收入中的成本为 85.97 元，主营业务收入利润率为 5.57%。

三、固定资产投资增速回落

全年固定资产投资（不含农户）551 590 亿元，比上年名义增长 10.0%，扣除价格因素实际增长 12.0%，实际增速比上年回落 2.9%。其中，国有控股投资 178 933 亿元，增长 10.9%；民间投资 354 007 亿元，增长 10.1%，占全部投资的比重为 64.2%。分产业看，第一产业投资 15 561 亿元，比上年增长 31.8%；第二产业投资 224 090 亿元，增长 8.0%；第三产业投资 311 939 亿元，增长 10.6%。从到位资金情况看，全年到位资金 573 789 亿元，比上年增长 7.7%。其中，国家预算资金增长 15.6%，国内贷款下降 5.8%，自筹资金增长 9.5%，利用外资下降 29.6%。全年新开工项目计划总投资 408 084 亿元，比上年增长 5.5%。从环比看，12 月份固定资产投资（不含农户）增长 0.68%。

全年全国房地产开发投资 95 979 亿元，比上年名义增长 1.0%（扣除价格因素实际增长 2.8%），其中住宅投资增长 0.4%。房屋新开工面积 154 454 万平方米，比上年下降 14.0%，其中住宅新开工面积下降 14.6%。全国商品房销售面积 128 495 万平方米，比上年增长 6.5%，其中住宅销售面积增长 6.9%。全国商品房销售额 87 281 亿元，比上年增长 14.4%，其中住宅销售额增长 16.6%。房地产开发企业土地购置面积 22 811 万平方米，比上年下降 31.7%。12 月末，全国商品房待售面积 71 853 万平方米，比上年末增长 15.6%。全年房地产开发企业到位资金 125 203 亿元，比上年增长 2.6%。

四、市场销售较快增长

全年社会消费品零售总额 300 931 亿元，比上年名义增长 10.7%（扣除价格因素实际增长 10.6%）。其中，限额以上单位消费品零售额 142 558 亿元，增长 7.8%。按经营单位所在地

分，城镇消费品零售额 258 999 亿元，比上年增长 10.5%，乡村消费品零售额 41 932 亿元，增长 11.8%。按消费形态分，餐饮收入 32 310 亿元，比上年增长 11.7%，商品零售 26 8621 亿元，增长 10.6%，其中限额以上单位商品零售 133 891 亿元，增长 7.9%。12 月份，社会消费品零售总额同比名义增长 11.1%（扣除价格因素实际增长 10.7%），环比增长 0.82%。

全年全国网上零售额 38 773 亿元，比上年增长 33.3%。其中，实物商品网上零售额 32 424 亿元，增长 31.6%，占社会消费品零售总额的比重为 10.8%；非实物商品网上零售额 6 349 亿元，增长 42.4%。

五、进出口同比下降

全年进出口总额 245 849 亿元，比上年下降 7.0%。其中，出口 141 357 亿元，下降 1.8%；进口 104 492 亿元，下降 13.2%。进出口相抵，顺差 36 865 亿元。12 月份，进出口总额 24 757 亿元，同比下降 0.5%。其中，出口 14 289 亿元，增长 2.3%；进口 10 468 亿元，下降 4.0%。

六、居民消费价格温和上涨

全年居民消费价格比上年上涨 1.4%。其中，城市上涨 1.5%，农村上涨 1.3%。分类别看，食品价格同比上涨 2.3%，烟酒及日用品上涨 2.1%，衣着上涨 2.7%，家庭设备用品及维修服务上涨 1.0%，医疗保健和个人用品上涨 2.0%，交通和通信下降 1.7%，娱乐教育文化用品及服务上涨 1.4%，居住上涨 0.7%。在食品价格中，粮食价格上涨 2.0%，油脂价格下降 3.2%，猪肉价格上涨 9.5%，鲜菜价格上涨 7.4%。12 月份，居民消费价格同比上涨 1.6%，环比上涨 0.5%。全年工业生产者出厂价格比上年下降 5.2%，12 月份同比下降 5.9%，环比下降 0.6%。全年工业生产者购进价格比上年下降 6.1%，12 月份同比下降 6.8%，环比下降 0.7%。

七、居民收入稳定增长

全年全国居民人均可支配收入 21 966 元，比上年名义增长 8.9%，扣除价格因素实际增长 7.4%。按常住地分，城镇居民人均可支配收入 31 195 元，比上年增长 8.2%，扣除价格因素实际增长 6.6%；农村居民人均可支配收入 11 422 元，比上年增长 8.9%，扣除价格因素实际增长 7.5%。城乡居民人均收入倍差 2.73，比上年缩小 0.02。全国居民人均可支配收入中位数 19 281 元，比上年名义增长 9.7%。按全国居民五等份收入分组，低收入组人均可支配收入 5 221 元，中等偏下收入组人均可支配收入 11 894 元，中等收入组人均可支配收入 19 320 元，中等偏上收入组人均可支配收入 29 438 元，高收入组人均可支配收入 54 544 元。2015 年全国居民收入基尼系数为 0.462。全年农民工总量 27 747 万人，比上年增加 352 万人，增长 1.3%，其中，本地农民工 10 863 万人，增长 2.7%，外出农民工 16 884 万人，增长 0.4%。农民工月均收入水平 3 072 元，比上年增长 7.2%。

八、经济结构优化升级

产业结构继续优化。全年第三产业增加值占国内生产总值的比重为 50.5%，比上年提高 2.4%，高于第二产业 10.0%。需求结构进一步改善。全年最终消费支出对国内生产总值增长的贡献率为 66.4%，比上年提高 15.4%。区域结构协调性增强。中、西部地区规模以上工业增加值比上年分别增长 7.6%和 7.8%，分别快于东部地区 0.9%和 1.1%；中部地区固定资产投资（不含农户）增长 15.7%，快于东部地区 3.0%。节能降耗继续取得新进展。全年单位国内生产总值能耗比上年下降 5.6%。

九、货币信贷平稳增长

12 月末，广义货币（M2）余额 139.23 万亿元，比上年末增长 13.3%，狭义货币（M1）余额 40.10 万亿元，增长 15.2%，流通中货币（M0）余额 6.32 万亿元，增长 4.9%。12 月末，人民币贷款余额 93.95 万亿元，人民币存款余额 135.70 万亿元。全年新增人民币贷款 11.72 万亿元，比上年多增 1.81 万亿元，新增人民币存款 14.97 万亿元，比上年多增 1.94 万亿元。全年社会融资规模增量为 15.41 万亿元。

十、人口就业总体稳定

年末中国大陆总人口（包括 31 个省、自治区、直辖市和中国人民解放军现役军人，不包括香港、澳门特别行政区和台湾省及海外华侨人数）137 462 万人，比上年末增加 680 万人。全年出生人口 1 655 万人，人口出生率为 12.07‰，死亡人口 975 万人，人口死亡率为 7.11‰，人口自然增长率为 4.96‰，比上年下降 0.25%。从性别结构看，男性人口 70 414 万人，女性人口 67 048 万人，总人口性别比为 105.02（以女性为 100），出生人口性别比为 113.51。从年龄构成看，16 周岁以上至 60 周岁以下（不含 60 周岁）的劳动年龄人口 91 096 万人，比上年末减少 487 万人，占总人口的比重为 66.3%；60 周岁及以上人口 22 200 万人，占总人口的 16.1%；65 周岁及以上人口 14 386 万人，占总人口的 10.5%。从城乡结构看，城镇常住人口 77 116 万人，比上年末增加 2 200 万人，乡村常住人口 60 346 万人，减少 1 520 万人，城镇人口占总人口比重为 56.1%。全国居住地和户口登记地不在同一个乡镇街道且离开户口登记地半年以上的人口（即人户分离人口）2.94 亿人，比上年末减少 377 万人，其中流动人口为 2.47 亿人，比上年末减少 568 万人。年末全国就业人员 77 451 万人，其中城镇就业人员 40 410 万人。

总的来看，2015 年国民经济仍运行在合理区间，经济结构进一步优化，转型升级进一步加快，新兴动力进一步积聚，人民生活进一步改善。但也要看到，国际环境仍然错综复杂，国内结构调整转型升级正处在爬坡过坎的关键阶段，全面深化改革任务艰巨。下一步，要认真学习贯彻党的十八大和十八届三、四、五中全会精神和中央经济工作会议精神，深入贯彻落实党中央、国务院决策部署，坚持宏观政策要稳、产业政策要准、微观政策要活、改革政策要实、社会政策要托底的总体思路，着力加强供给侧结构性改革，推动国民经济保持中高速、迈向中高端，实现中高收入水平。

附注：

①为进一步推动我国季度国内生产总值（GDP）核算与国际接轨，国家统计局从 2015 年三季度起实施季度 GDP 核算改革，采用分季方式核算 GDP 数据，并公布相关核算结果。

②国内生产总值、规模以上工业增加值及其分类项目增长速度按可比价计算，为实际增长速度；其他指标除特殊说明外，按现价计算，为名义增长速度。

③根据季节调整模型自动修正结果，对近一年来各期国内生产总值、规模以上工业增加值、固定资产投资（不含农户）、社会消费品零售总额环比增速进行修订。修订结果及 2015 年四季度 GDP 环比数据、2015 年 12 月份其他指标环比数据如下。

2014 年及 2015 年各季度 GDP 环比增速分别为 1.6%、1.8%、1.8%、1.7%、1.3%、1.9%、1.8%和 1.6%。

资料来源：http://www.chinairn.com/news/20160119/102742315.shtml

参 考 文 献

[1] 延静. 调查技能与分析[M]. 北京：清华大学出版社，2006.
[2] 李洁. 统计学[M]. 长春：东北师范大学出版社，2008.
[3] 陈在余，陶应虎. 统计学原理与实务[M]. 北京：清华大学出版社，2009.
[4] 段雪妍，肖智明. 统计学原理[M]. 上海：上海财经大学出版社，2009.
[5] 谢景文，吴小平. 统计学原理[M]. 北京：北京理工大学出版社，2010.
[6] 卜小玲，李洁. 统计学原理与实务[M]. 北京：清华大学出版社，2010.
[7] 邢西治. 统计学原理学习指导与习题解析[M]. 南京：南京大学出版社，2010.
[8] 韩兆洲. 统计学原理（第七版）[M]. 广州：暨南大学出版社，2011.
[9] 王云峰. 统计学原理：理论与方法[M]. 上海：复旦大学出版社，2010.
[10] 贾俊平，郝静. 统计学案例与分析[M]. 北京：中国人民大学出版社，2010.
[11] 袁卫，刘超. 统计学：思想、方法与应用[M]. 北京：中国人民大学出版社，2011.
[12] 杨立生. 统计学原理[M]. 北京: 北京大学出版社，2011.
[13] 林侠，郑小丽. 统计学原理与实务[M]. 北京：北京师范大学出版社，2011.
[14] 胡晓晖，张文喜. 应用统计与 Excel 运用[M]. 上海：上海财经大学出版社，2011.
[15] 张宏. 统计学原理[M]. 北京：中国铁道出版社，2012.
[16] 徐静霞. 统计学原理与实务[M]. 北京：北京大学出版社，2012.
[17] 吴思莹，邢小博. 统计学原理[M]. 北京：北京大学出版社，2013.
[18] 李博达，阎薇，肖旭. 统计学原理[M]. 北京：北京交通大学出版社，2014.
[19] 徐晓岭，王磊. 统计学[M]. 北京：人民邮电出版社，2015.
[20] 卢淑华. 社会统计学（第 4 版）[M]. 北京：北京大学出版社，2015.
[21] 孙静娟. 统计学[M]. 北京：清华大学出版社，2015.
[22] 贾俊平. 统计学（第六版）[M]. 北京：中国人民大学出版社，2016.
[23] 北京统计信息网. http://www.bjstats.gov.cn/.
[24] 中华人民共和国国家统计局网站. http://www.stats.gov.cn/.
[25] 中华人民共和国商务部. http://www.mofcom.gov.cn/.
[26] 中华人民共和国海关总署. http://www.customs.gov.cn/publish/portal0/.
[27] 中国物流与采购联合会. http://www.chinawuliu.com.cn/.
[28] 中国人口信息网. http://old.cpwf.org.cn/.
[29] 中国农业普查网. http://www.stats.gov.cn/zgnypc/.
[30] 中国统计信息网. http://www.tjcn.org/.
[31] 中国行业研究网. http://market.chinairn.com/diaochazhixing03.html.
[32] 中国大数据. http://www.thebigdata.cn/.
[33] 中国统计网. http://www.itongji.cn/cms/article/index.

反侵权盗版声明

电子工业出版社依法对本作品享有专有出版权。任何未经权利人书面许可，复制、销售或通过信息网络传播本作品的行为；歪曲、篡改、剽窃本作品的行为，均违反《中华人民共和国著作权法》，其行为人应承担相应的民事责任和行政责任，构成犯罪的，将被依法追究刑事责任。

为了维护市场秩序，保护权利人的合法权益，我社将依法查处和打击侵权盗版的单位和个人。欢迎社会各界人士积极举报侵权盗版行为，本社将奖励举报有功人员，并保证举报人的信息不被泄露。

举报电话：（010）88254396；（010）88258888
传　　真：（010）88254397
E-mail：　dbqq@phei.com.cn
通信地址：北京市万寿路 173 信箱
　　电子工业出版社总编办公室
邮　　编：100036